F. GANTMACHER

Lectures in Analytical Mechanics

TRANSLATED FROM THE RUSSIAN
BY GEORGE YANKOVSKY

MIR PUBLISHERS MOSCOW

First published 1970
Second printing 1975

This book has been recommended for publication by Indian specialists under the programme of the Joint Indo-Soviet Board to make the best Soviet textbooks available for Indian students

На английском языке

© English translation, Mir Publishers, 1975

Contents

Preface 7

Chapter 1. THE DIFFERENTIAL EQUATIONS OF MOTION OF AN ARBITRARY SYSTEM OF PARTICLES 9

1. Free and Constrained Systems. Constraints and Their Classification 9
2. Possible and Virtual Displacements. Ideal Constraints . . . 12
3. The General Equation of Dynamics. Lagrange's Equations of the First Kind 20
4. The Principle of Virtual Displacements. D'Alembert's Principle 25
5. Holonomic Systems. Independent Coordinates. Generalized Forces 33
6. Lagrange's Equations of the Second Kind in Independent Coordinates 40
7. Investigating Lagrange's Equations 45
8. Theorem on Variation of Total Energy. Potential, Gyroscopic and Dissipative Forces 48
9. Electromechanical Analogies 54
10. Appell's Equations for Nonholonomic Systems. Pseudocoordinates 57

Chapter 2. THE EQUATIONS OF MOTION IN A POTENTIAL FIELD 66

11. Lagrange's Equations for Potential Forces. The Generalized Potential. Nonnatural Systems 66
12. Canonical Equations of Hamilton 71
13. Routh's Equations 78
14. Cyclic Coordinates 80
15. The Poisson Bracket 83

Chapter 3. VARIATIONAL PRINCIPLES AND INTEGRAL INVARIANTS 88

16. Hamilton's Principle 88
17. Second Form of Hamilton's Principle 96
18. The Basic Integral Invariant of Mechanics (Poincaré-Cartan Integral Invariant) 98
19. A Hydrodynamical Interpretation of the Basic Integral Invariant. The Theorems of Thomson and Helmholtz on Circulation and Vortices 105
20. Generalized Conservative Systems. Whittaker's Equations. Jacobi's Equations. The Maupertuis-Lagrange Principle of Least Action 111
21. Inertial Motion. Relation to Geodesic Lines in the Arbitrary Motion of a Conservative System 117
22. The Universal Integral Invariant of Poincaré. Lee Hwa-Chung's Theorem 119
23. Invariance of Volume in the Phase Space. Liouville's Theorem 125

Chapter 4. CANONICAL TRANSFORMATIONS AND THE HAMILTON-JACOBI EQUATION 128

24. Canonical Transformations 128
25. Free Canonical Transformations 132
26. The Hamilton-Jacobi Equation 135
27. Method of Separation of Variables. Examples 141

28. Applying Canonical Transformations to Perturbation Theory 150
29. The Structure of an Arbitrary Canonical Transformation . . 152
30. Testing the Canonical Character of a Transformation. The Lagrange Brackets 158
31. The Simplicial Nature of the Jacobian Matrix of a Canonical Transformation 161
32. Invariance of the Poisson Brackets in a Canonical Transformation 163

Chapter 5. STABILITY OF EQUILIBRIUM AND THE MOTIONS OF A SYSTEM 166

33. Lagrange's Theorem on the Stability of an Equilibrium Position 166
34. Criteria of Instability of an Equilibrium Position. Theorems of Lyapunov and Chetayev 173
35. Asymptotic Stability of an Equilibrium Position. Dissipative Systems 176
36. Conditional Stability. General Statement of the Problem. Stability of Motion or of an Arbitrary Process. Lyapunov's Theorem 182
37. Stability of Linear Systems 189
38. Stability in Linear Approximation 193
39. Criteria of Asymptotic Stability of Linear Systems 197

Chapter 6. SMALL OSCILLATIONS 202

40. Small Oscillations of a Conservative System 202
41. Normal Coordinates 211
42. The Effect of Periodic External Forces on the Oscillations of a Conservative System 214
43. Extremal Properties of Frequencies of a Conservative System. Rayleigh's Theorem on Frequency Variation with Change in Inertia and Rigidity of the System. Superimposition of Constraints 216
44. Small Oscillations of Elastic Systems 222
45. Small Oscillations of a Scleronomic System under the Action of Forces Not Explicitly Dependent on the Time 228
46. Rayleigh's Dissipative Function. The Effect of Small Dissipative Forces on the Oscillations of a Conservative System 231
47. The Effect of a Time-Dependent External Force on Small Oscillations of a Scleronomic System. The Amplitude-Phase Characteristic 235

Chapter 7. SYSTEMS WITH CYCLIC COORDINATES 242

48. Reduced System. The Routh Potential. Hidden Motions. Hertz' Conception of the Kinetic Origin of Potential Energy 242
49. Stability of Stationary Motions 252

References 259

Name Index 260

Subject Index 261

Preface

In the literature on mechanics there is no single generally accepted interpretation of the term "analytical mechanics". Some writers identify analytical mechanics with theoretical mechanics.* Others maintain that an exposition in generalized coordinates constitutes the distinguishing feature of analytical mechanics. A third point of view — the one adhered to by the author in calling this book "Lectures in Analytical Mechanics"—consists in the fact that analytical mechanics is characterized both by a specific system of presentation and also by a definite range of problems investigated.

The characteristic feature of the system of presentation of analytical mechanics is that general principles (differential or integral) serve as the foundation; then the basic differential equations of motion are derived from these principles analytically. The basic content of analytical mechanics consists in describing the general principles of mechanics, deriving from them the fundamental differential equations of motion, and investigating the equations obtained and methods of integrating them.

Analytical mechanics constitutes a part of the course of theoretical mechanics that comes within the syllabus of mechanico-mathematical, physical and physical-engineering departments of universities and pedagogical institutes. Yet the general syllabus in theoretical mechanics of technological institutes either does not contain analytical mechanics at all or preserves only elements of it, though modern technology poses problems that lie beyond the fundamental course of theoretical mechanics as given in the traditional divisions of "statics", "kinematics", and "the dynamics of a particle and a system". Research engineers working in diverse areas of modern technology must also master the general methods of analytical mechanics which yield a universal analytical apparatus for investigating complex problems that refer not only to purely mechanical, but also to electrical and electromechanical phenomena.

---

* For example, the well-known courses of theoretical mechanics of G. K. Suslov and Ch. de la Vallée-Poussin were called courses of analytical mechanics by the authors.

The present text does not claim to cover fully the material of analytical mechanics. It developed out of a course of lectures given by the author over a period of six years in the Moscow Physico-Technical Institute. This circumstance determined the choice of material and the manner of its presentation.

The course of analytical mechanics is a foundation supporting such divisions of theoretical physics as quantum mechanics, the special and general theories of relativity, and so forth. For this reason, a detailed presentation is given of variational principles and the integral invariants of mechanics, canonical transformations, the Hamilton-Jacobi equation, and systems with cyclic (ignorable) coordinates (Chapters 2, 3, 4 and 7). Following the ideas of Poincaré and Cartan, the author takes the integral invariants of mechanics as the basis of presentation. Here they do not represent an embellishment of the theory but its actual workaday machinery.

The technical applications are associated with a consideration of constrained systems, which are studied in detail in Chapter 1. In a special section of that chapter, which is devoted to electromechanical analogies, the possibility is investigated of extending the analytical methods of mechanics to electrical and electromechanical systems. In Chapters 5 and 6 are given applications of analytical mechanics to Lyapunov's theory of stability and the theory of oscillations. Elements of modern frequency methods are given along with the classical problems in the theory of linear oscillations. Problems in the dynamics of rigid bodies are taken up in individual examples.

It is assumed the reader is acquainted with the general fundamentals of theoretical mechanics and higher mathematics. The text is designed for undergraduate and graduate students of mechanico-mathematical, physical and physical-engineering departments of universities, and also for research engineers and other specialists who feel a need to extend and deepen their knowledge in the field of mechanics.

CHAPTER 1

The Differential Equations of Motion of an Arbitrary System of Particles

1. Free and Constrained Systems. Constraints and Their Classification

The motion is studied of a system of particles P_ν ($\nu = 1, \ldots, N$) relative to some inertial (Galilean) system of coordinates. Imposed on the positions and velocities of the particles of the system are restrictions of a geometrical or kinematical nature, called constraints. Systems with such constraints are known as constrained systems in contradistinction to free systems, in which such constraints are absent.

Analytically a constraint is expressed by the equation

$$f(t, \boldsymbol{r}_\nu, \dot{\boldsymbol{r}}_\nu) = 0^{*} \tag{1}$$

where the left-hand side includes the time t, the radius vectors $\boldsymbol{r}_\nu$ and the velocities $\boldsymbol{v}_\nu = \dot{\boldsymbol{r}}_\nu$ of all points P_ν of the system ($\nu = 1, \ldots, N$). In the special case when the velocities $\dot{\boldsymbol{r}}_\nu$ do not enter into the constraint equation (1), the constraint is termed *finite* or *geometric*. Analytically, it is written as

$$f(t, \boldsymbol{r}_\nu) = 0 \tag{2}$$

In the general case, the constraint (1) is called *differential* or *kinematical*. From now on we shall confine our consideration solely to such differential constraints whose equations contain the velocities of the particles in linear form:

$$\sum_{\nu=1}^{N} \boldsymbol{l}_\nu \dot{\boldsymbol{r}}_\nu + D = 0 \tag{3}$$

---

[*] A dot over a letter denotes differentiation of the given quantity with respect to time. All radius vectors are constructed from a single pole that is stationary in the given system of coordinates. Also, $f(t, \boldsymbol{r}_\nu, \dot{\boldsymbol{r}}_\nu)$ is an abridged notation for the function $f(t, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_N, \dot{\boldsymbol{r}}_1, \ldots, \dot{\boldsymbol{r}}_N)$. Such abbreviated notation will be used throughout this text. If x_ν, y_ν, and z_ν are the Cartesian coordinates of a point P_ν in the given system of coordinates ($\nu = 1, \ldots, N$) then the function f may be regarded as a function of $6N + 1$ scalar arguments t, x_ν, y_ν, z_ν, $\dot{x}_\nu$, $\dot{y}_\nu$, and $\dot{z}_\nu$ ($\nu = 1, \ldots, N$). The function f and all other functions that will be encountered in the sequel, are assumed (unless otherwise stated) to be continuous together with those of their derivatives which occur in appropriate parts of the text.

where $\boldsymbol{l}_\nu \dot{\boldsymbol{r}}_\nu$ is the scalar product of the vectors $\boldsymbol{l}_\nu$ and $\dot{\boldsymbol{r}}_\nu$, and the vectors $\boldsymbol{l}_\nu$ and the scalar D are specified functions of t and of all $\boldsymbol{r}_\mu$ (μ, $\nu = 1, \ldots, N$). It is assumed here that the vectors $\boldsymbol{l}_\nu$ cannot all vanish at the same time.

Given a finite constraint of type (2), a system cannot occupy an arbitrary position in space at every given instant of time. Finite constraints impose restrictions to possible positions of the system at time t. But with a differential constraint alone, the system may occupy any arbitrary position in space at any time t. However, in this position the velocities of the particles of the system cannot any longer be arbitrary, since the differential constraint imposes restrictions on these velocities.

Each finite constraint of type (2) implies, as a consequence, a differential constraint whose equation is obtained by termwise differentiation of equation (2):

$$\sum_{\nu=1}^{N} \frac{\partial f}{\partial \boldsymbol{r}_\nu} \dot{\boldsymbol{r}}_\nu + \frac{\partial f}{\partial t} = 0 \tag{4}$$

where $\frac{\partial f}{\partial \boldsymbol{r}_\nu} = \operatorname{grad}_\nu f$ $(\nu = 1, \ldots, N)$.* But such a differential constraint is not equivalent to the finite constraint (2). It is equivalent to the finite constraint

$$f(t, \boldsymbol{r}_\nu) = c \tag{5}$$

where c is an arbitrary constant. For this reason, the finite constraint (4) is called *integrable*.

It will be noted that in rectangular Cartesian coordinates, the constraint equations (1) to (4) are written as:

$$f(t, x_\nu, y_\nu, z_\nu, \dot{x}_\nu, \dot{y}_\nu, \dot{z}_\nu) = 0 \tag{1'}$$

$$f(t, x_\nu, y_\nu, z_\nu) = 0 \tag{2'}$$

$$\sum_{\nu=1}^{N} (A_\nu \dot{x}_\nu + B_\nu \dot{y}_\nu + C_\nu \dot{z}_\nu) + D = 0^{**} \tag{3'}$$

$$\sum_{\nu=1}^{N} \left(\frac{\partial f}{\partial x_\nu} \dot{x}_\nu + \frac{\partial f}{\partial y_\nu} \dot{y}_\nu + \frac{\partial f}{\partial z_\nu} \dot{z}_\nu \right) + \frac{\partial f}{\partial t} = 0 \tag{4'}$$

---

* If $\boldsymbol{r}_\nu = x_\nu \boldsymbol{i} + y_\nu \boldsymbol{j} + z_\nu \boldsymbol{k}$, where $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ are mutually orthogonal unit vectors of the coordinate axes, then

$$\frac{\partial f}{\partial \boldsymbol{r}_\nu} = \frac{\partial f}{\partial x_\nu} \boldsymbol{i} + \frac{\partial f}{\partial y_\nu} \boldsymbol{j} + \frac{\partial f}{\partial z_\nu} \boldsymbol{k} \; (\nu = 1, \ldots, N)$$

** A_ν, B_ν, and C_ν $(\nu = 1, \ldots, N)$ are scalar functions of t, x_1, y_1, z_1, ... x_N, y_N, z_N.

The finite constraint (2) or (2′) is termed *stationary* if t is not expressed explicitly in the constraint equation, i.e. if $\frac{\partial f}{\partial t} = 0$. In this case, the left-hand side of the equation of differential constraint (4) is linear and homogeneous in the velocities. By analogy, the differential constraint (3) or (3′) is termed *stationary* if $D = 0$ and if the vectors $\boldsymbol{l}_\nu$ in equation (3) [and, respectively, the coefficients A_ν, B_ν, and C_ν in equation (3′)] are not explicit functions of t.

A system of particles is called *holonomic* if the particles of the system are not subjected to differential nonintegrable constraints. Thus, a holonomic system is any free system of particles and also any constrained system with finite or differential but integrable constraints. All constraints in a holonomic system may be written in closed form.

A system is called *nonholonomic*[*] if there are differential nonintegrable constraints.

A system is termed *scleronomic* if only stationary constraints are imposed, otherwise it is *rheonomic*.

Example 1. A particle is constrained to move over a surface. Let the equation of this surface be given in the form

$$f(\boldsymbol{r}) = 0 \tag{6}$$

or

$$f(x, y, z) = 0. \tag{6'}$$

This is a finite stationary constraint.

If the surface is moving or undergoing deformation, the time t enters the equation of the surface explicitly:

$$f(t, \boldsymbol{r}) = 0 \tag{7}$$

or

$$f(t, x, y, z) = 0. \tag{7'}$$

In this case, the constraint is finite but nonstationary.

Example 2. Two particles are connected by a rod of constant length l. Here the constraint equation is of the form

$$(\boldsymbol{r}_1 - \boldsymbol{r}_2)^2 - l^2 = 0 \tag{8}$$

or

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - l^2 = 0. \tag{8'}$$

This is a holonomic scleronomic system.

Let us note that a rigid body may be regarded as a system of particles equidistant from one another, that is, subjected to constraints of type (8). From this point of view, a free rigid body is a special case of a constrained holonomic scleronomic system of particles.

Example 3. Two particles are connected by a rod of variable length $l = f(t)$. The constraint equation is written thus:

$$(\boldsymbol{r}_1 - \boldsymbol{r}_2)^2 - f^2(t) = 0 \tag{9}$$

---

[*] Nonintegrable differential constraints are themselves frequently called nonholonomic. Sometimes integrable differential constraints are termed semiholonomic.

or

$$(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2-f^2(t)=0. \tag{9'}$$

This is a holonomic rheonomic system.

Example 4. Two particles in a plane are connected by a rod of constant length l and are constrained to move in such a manner that the velocity of the middle of the rod is in the direction of the rod (the motion of a skate on a plane). The constraint equations are written as follows:

$$\left.\begin{aligned} &z_1=0, \quad z_2=0 \\ &(x_1-x_2)^2+(y_1-y_2)^2-l^2=0 \\ &\frac{\dot{x}_1+\dot{x}_2}{x_1-x_2}=\frac{\dot{y}_1+\dot{y}_2}{y_1-y_2} \end{aligned}\right\} \tag{10}$$

This is a nonholonomic system because the last equation in (10) defines a nonintegrable differential constraint.

In addition to constraints of type (1), which are called *bilateral* constraints, mechanics also studies *unilateral* constraints, which are written in the form of an inequality:

$$f(t, \boldsymbol{r}_\nu, \dot{\boldsymbol{r}}_\nu) \geqslant 0 \tag{11}$$

An example is the case of two particles connected by a thread of length l. This constraint is expressed by the inequality

$$l^2-(\boldsymbol{r}_1-\boldsymbol{r}_2)^2 \geqslant 0 \tag{12}$$

If in condition (11) we have an equal sign, it is said that the *constraint is taut*.

The motion of a system on which a unilateral constraint is imposed may be divided into portions so that in certain portions the constraint is taut and the motion occurs as if the constraint were bilateral, and in other portions the constraint is not taut and the motion occurs as if there were no such constraint. In other words, in certain portions a unilateral constraint is either replaced by a bilateral constraint or is eliminated altogether. For this reason we shall henceforth consider only bilateral constraints.

2. Possible and Virtual Displacements. Ideal Constraints

Let there be imposed, on a material system, d finite constraints

$$f_\alpha(t, \boldsymbol{r}_\nu)=0 \quad (\alpha=1, \ldots, d) \tag{1}$$

and g differential constraints*

$$\sum_{\nu=1}^{N} \boldsymbol{l}_{\beta\nu}\boldsymbol{v}_\nu+D_\beta=0 \quad (\beta=1, \ldots, g) \tag{2}$$

---

* In equations of differential constraints, in place of $\dot{\boldsymbol{r}}_\nu$ we write $\boldsymbol{v}_\nu$.

We replace the finite constraints by the differential constraints that follow from them:

$$\sum_{\nu=1}^{N} \frac{\partial f_\alpha}{\partial \boldsymbol{r}_\nu} \boldsymbol{v}_\nu + \frac{\partial f_\alpha}{\partial t} = 0 \quad (\alpha = 1, \ldots, d) \tag{3}$$

The system of vectors $\boldsymbol{v}_\nu$ will be called *possible velocities* for a certain instant of time t and for a certain possible (at that instant) position of the system if the vectors $\boldsymbol{v}_\nu$ satisfy $d+g$ linear equations (2) and (3).

Thus, possible velocities are velocities permitted by the constraints. For every possible position of the system at time t there exists an infinity of systems of possible velocities. One of these systems of velocities is realized in the actual motion of the system at time t.

We shall call the system of infinitesimal displacements

$$d\boldsymbol{r}_\nu = \boldsymbol{v}_\nu \, dt \quad (\nu = 1, \ldots, N) \tag{4}$$

where $\boldsymbol{v}_\nu$ $(\nu = 1, \ldots, N)$ are the possible velocities, *possible infinitesimal displacements* or, for brevity, simply *possible displacements*. Multiplying equations (2) and (3) termwise by dt, we get equations that determine the possible displacements:

$$\left.\begin{aligned} &\sum_{\nu=1}^{N} \frac{\partial f_\alpha}{\partial \boldsymbol{r}_\nu} d\boldsymbol{r}_\nu + \frac{\partial f_\alpha}{\partial t} dt = 0 \quad (\alpha = 1, \ldots, d) \\ &\sum_{\nu=1}^{N} \boldsymbol{l}_{\beta\nu} \, d\boldsymbol{r}_\nu + D_\beta \, dt = 0 \quad (\beta = 1, \ldots, g) \end{aligned}\right\} \tag{5}$$

Let us take two systems of possible displacements at one and the same instant of time and for one and the same position of the system:

$$d\boldsymbol{r}_\nu = \boldsymbol{v}_\nu \, dt, \text{ and } d'\boldsymbol{r}_\nu = \boldsymbol{v}'_\nu \, dt \quad (\nu = 1, \ldots, N).$$

Both $d\boldsymbol{r}_\nu$ and $d'\boldsymbol{r}_\nu$ satisfy the equations (5), while the differences

$$\delta \boldsymbol{r}_\nu = d'\boldsymbol{r}_\nu - d\boldsymbol{r}_\nu \quad (\nu = 1, \ldots, N) \tag{6}$$

satisfy the homogeneous relations

$$\left.\begin{aligned} &\sum_{\nu=1}^{N} \frac{\partial f_\alpha}{\partial \boldsymbol{r}_\nu} \delta \boldsymbol{r}_\nu = 0 \quad (\alpha = 1, \ldots, d) \\ &\sum_{\nu=1}^{N} \boldsymbol{l}_{\beta\nu} \delta \boldsymbol{r}_\nu = 0 \quad (\beta = 1, \ldots, g) \end{aligned}\right\} \tag{7}$$

The differences $\delta \boldsymbol{r}_\nu = d'\boldsymbol{r}_\nu - d\boldsymbol{r}_\nu$ will be called *virtual displacements*. Any system of vectors $\delta \boldsymbol{r}_\nu$ satisfying equations (7) is a system of virtual displacements. The equations (7) for virtual displacements differ from equations (5), which define possible displacements.

in the absence of the terms $\frac{\partial f_\alpha}{\partial t}\,dt$ and $D_\beta\,dt$. It is said for this reason that *virtual displacements coincide with possible displacements in the case of 'frozen' constraints.*

Indeed, in 'freezing', the time t that enters into the equations of finite constraints becomes fixed, that is the constraint congeals, as it were, in the configuration that it had at time t. Then the terms $\frac{\partial f_\alpha}{\partial t}\,dt$ do not appear when differentiating the functions f_α, and the first d equations (5) coincide with the corresponding equations (7). For a differential constraint, 'freezing' signifies imposing on it a stationary character, i.e. the elimination of D_β on the left-hand side of the constraint equation and the fixing of t which enters explicitly in the coefficients $l_{\beta\nu}$. Then the last g equations (5) will also coincide with the corresponding equations (7).

We can say that virtual displacements are displacements of points of a system from one possible position of the system at time t to another infinitely close possible (for the same instant of time t) position of the system.

In the case of stationary constraints, virtual displacements coincide with possible displacements.

Example 1. A particle is in motion on a fixed surface (Fig. 1).

In this case, any vector $\boldsymbol{v}$ constructed from the point P and tangent to the surface at this point will constitute a possible velocity. The corresponding possible displacement $d\boldsymbol{r} = \boldsymbol{v}\,dt$ likewise lies in the plane tangent to the surface at the point P. The difference $\delta\boldsymbol{r} = d'\boldsymbol{r} - d\boldsymbol{r}$ of the two tangent vectors is in turn a vector tangent to the surface at this point. Thus, any vector constructed from P and lying in the tangent plane may be regarded as a certain $d\boldsymbol{r}$ and as a certain $\delta\boldsymbol{r}$. Here the constraint is stationary and the virtual displacements coincide with the possible displacements.

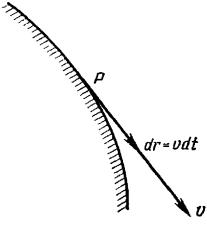

Fig. 1.

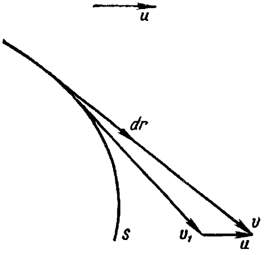

Fig. 2.

Example 2. The constraint is a surface S which is itself in motion (as a rigid body) with a certain velocity $\boldsymbol{u}$ relative to the original system of coordinates (Fig. 2).

In this case, the possible velocity $\boldsymbol{v}$ is obtained from an arbitrary vector $\boldsymbol{v}_1$ that is tangent to the surface by adding to it the velocity $\boldsymbol{u}$:

$$\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{u}$$

Therefore,

$$d\boldsymbol{r} = \boldsymbol{v}\, dt = \boldsymbol{v}_1\, dt + \boldsymbol{u}\, dt.$$

Similarly, for another possible displacement

$$d'\boldsymbol{r} = \boldsymbol{v}_1'\, dt + \boldsymbol{u}\, dt$$

the virtual displacement

$$\delta\boldsymbol{r} = d'\boldsymbol{r} - d\boldsymbol{r} = (\boldsymbol{v}_1' - \boldsymbol{v}_1)\, dt$$

is (unlike $d\boldsymbol{r}$) a vector lying in a plane tangent to the surface at the point P (Fig. 3). The vector $\delta\boldsymbol{r}$ is a possible displacement for the 'stopped' surface S.

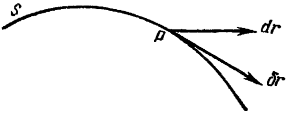

Fig. 3.

In Cartesian coordinates, the vector $\delta\boldsymbol{r}_\nu$ is characterized by three projections on the axes δx_ν, δy_ν, δz_ν $(\nu = 1, \ldots, N)$, and the equations (7) which define the virtual displacements may be written in the following form:

$$\left.\begin{aligned}&\sum_{\nu=1}^{N}\left(\frac{\partial f_\alpha}{\partial x_\nu}\delta x_\nu + \frac{\partial f_\alpha}{\partial y_\nu}\delta y_\nu + \frac{\partial f_\alpha}{\partial z_\nu}\delta z_\nu\right) = 0 \quad (\alpha = 1, \ldots, d)\\ &\sum_{\nu=1}^{N}(A_{\beta\nu}\delta x_\nu + B_{\beta\nu}\delta y_\nu + C_{\beta\nu}\delta z_\nu) = 0 \quad (\beta = 1, \ldots, g)\end{aligned}\right\} \tag{7'}$$

If these $d + g$ equations are independent, then among $3N$ virtual increments of the coordinates δx_ν, δy_ν, δz_ν there will be $n = 3N - d - g$ independent virtual increments, where n is called the *number of degrees of freedom* of the given system of particles.

Let the corresponding forces $\boldsymbol{F}_\nu$ $(\nu = 1, \ldots, N)$ be impressed at the points P_ν of the system.* If the constraints were absent, then by Newton's second law we would have the relations $m_\nu \boldsymbol{w}_\nu = \boldsymbol{F}_\nu$ $(\nu = 1, \ldots, N)$ between the masses m_ν, the accelerations $\boldsymbol{w}_\nu$ and the forces $\boldsymbol{F}_\nu$. Given constraints, the accelerations $\boldsymbol{w}_\nu = \frac{1}{m_\nu}\boldsymbol{F}_\nu$ (at a given instant of time t, in a given position of the particles of the system $\boldsymbol{r}_\nu$, and for given velocities $\boldsymbol{v}_\nu$) may prove incompatible with the constraints. Indeed, differentiating the equations (3) and (2) termwise with respect to the time, we get an analytical expression for the restrictions imposed by the constraints on the accelerations $\boldsymbol{w}_\nu$ of the particles of the system:*

---

* By $\boldsymbol{F}_\nu$ is understood the resultant of all forces applied directly to the particle P_ν $(\nu = 1, \ldots, N)$.

$$\left.\begin{aligned} &\sum_{\nu=1}^{N} \frac{\partial f_\alpha}{\partial \boldsymbol{r}_\nu}\boldsymbol{w}_\nu + \sum_{\nu=1}^{N}\left(\frac{d}{dt}\frac{\partial f_\alpha}{\partial \boldsymbol{r}_\nu}\right)\boldsymbol{v}_\nu + \frac{d}{dt}\frac{\partial f_\alpha}{\partial t} = 0 \quad (\alpha = 1, \ldots, d)\\ &\sum_{\nu=1}^{N} \boldsymbol{l}_{\beta\nu}\boldsymbol{w}_\nu + \sum_{\nu=1}^{N}\frac{d\boldsymbol{l}_{\beta\nu}}{dt}\boldsymbol{v}_\nu + \frac{dD_\beta}{dt} = 0 \quad (\beta = 1, \ldots, g) \end{aligned}\right\} \tag{8}$$

The accelerations $\boldsymbol{w}_\nu = \frac{1}{m_\nu}\boldsymbol{F}_\nu$ may not satisfy these relations. Then the materially effected constraints will act on the particles of the system P_ν with certain supplementary forces $\boldsymbol{R}_\nu$ ($\nu = 1, \ldots, N$); these forces are called the *reaction forces of the constraints.*** The reactions that arise are such that the accelerations determined from the equations

$$m_\nu \boldsymbol{w}_\nu = \boldsymbol{F}_\nu + \boldsymbol{R}_\nu \quad (\nu = 1, \ldots, N) \tag{9}$$

are already permitted by the constraints.

Unlike the reactions $\boldsymbol{R}_\nu$ ($\nu = 1, \ldots, N$), the preassigned forces $\boldsymbol{F}_\nu$ ($\nu = 1, \ldots, N$) are termed *effective forces*. Effective forces are ordinarily specified as known functions of the time, position and velocities of the particles of the system:***

$$\boldsymbol{F}_\nu = \boldsymbol{F}_\nu(t, \boldsymbol{r}_\nu, \boldsymbol{v}_\nu) \quad (\nu = 1, \ldots, N). \tag{10}$$

The basic problem of the dynamics of a constrained system consists in the following.

Given the effective forces $\boldsymbol{F}_\nu = \boldsymbol{F}_\nu(t, \boldsymbol{r}_\nu, \boldsymbol{v}_\nu)$ and the initial positions $\boldsymbol{r}_\nu^0$ and the initial velocities $\boldsymbol{v}_\nu^0$ of the particles of the system ($\nu = 1, \ldots, N$)—both are compatible with constraints—it is required to determine the motion of the system and the reactions of the constraints $\boldsymbol{R}_\nu$ ($\nu = 1, \ldots, N$).****

---

* The left sides in relations (8) are linearly dependent on the accelerations $\boldsymbol{w}_\nu$. As will readily be seen after differentiation, these left-hand sides are also dependent on t, $\boldsymbol{r}_\nu$, $\boldsymbol{v}_\nu$ ($\nu = 1, \ldots, N$).

** In the case of several constraints ($d + g > 1$), $\boldsymbol{R}_\nu$ is the resultant of all the reactions of constraints for the point P_ν ($\nu = 1, \ldots, N$).

*** In the general case, the right-hand sides of (10) depend on t and also on all $\boldsymbol{r}_\nu$ and $\boldsymbol{v}_\nu$ ($\nu = 1, \ldots, N$).

**** In the case of a free system, there is no problem of determining reactions and there only remains the problem of determining the motion of the system.

If nothing is known about the nature of the constraints except the defining equations (1) and (2) and, consequently, nothing is known about the reactions $\boldsymbol{R}_\nu$ produced by these constraints, then the above problem is indeterminate, since the number of scalar quantities x_ν, y_ν, z_ν, $R_{\nu x}$, $R_{\nu y}$, $R_{\nu z}$ that have to be determined is greater than the number of available scalar relations, i.e. the equations $m_\nu \ddot{x}_\nu = F_{\nu x} + R_{\nu x}$, $m_\nu \ddot{y}_\nu = F_{\nu y} + R_{\nu y}$, $m_\nu \ddot{z}_\nu = F_{\nu z} + R_{\nu z}$ and the constraint equations (1) and (2) $[6N > 3N + d + g]$.

For the basic problem of dynamics to become determinate, it is necessary to have some kind of additional $6N - (3N + d + g) = 3N - d - g = n$ independent relations between the sought-for quantities. These relations can be obtained if we confine ourselves to the important class of *ideal constraints*.

Constraints are termed ideal if the sum of the works of the reactions of these constraints on any virtual displacements is equal to zero, i.e. if

$$\sum_{\nu=1}^{N} \boldsymbol{R}_\nu \delta \boldsymbol{r}_\nu = 0. \tag{11}$$

This equation may be rewritten in expanded form:

$$\sum_{\nu=1}^{N} (R_{\nu x}\delta x_\nu + R_{\nu y}\delta y_\nu + R_{\nu z}\delta z_\nu) = 0. \tag{11'}$$

Among the $3N$ quantities δx_ν, δy_ν, δz_ν, there are n independent ones ($n = 3N - d - g$ is the number of degrees of freedom of the given system). It is therefore possible in (11') to express $3N - n$ dependent increments δx_ν, δy_ν, δz_ν in terms of n independent increments and equate to zero the coefficients of these independent increments. We then obtain the n relations still lacking and needed to make determinate the basic problem of the dynamics of a constrained system.

After considering the following examples of ideal constraints it will be clear how natural and important practically is the class of constraints that we have isolated.

Example 1. A particle is constrained to move on a stationary smooth surface (Fig. 4).

In this case, any possible displacement $d\boldsymbol{r}$, like any virtual displacement $\delta\boldsymbol{r}$, lies in a plane that is tangent to the surface at the point P, and the reaction of the smooth surface is directed normally to the surface at this point. For this reason, we always have

$$\boldsymbol{R}\, d\boldsymbol{r} = 0 \text{ or } \boldsymbol{R}\delta\boldsymbol{r} = 0$$

Example 2. A particle is constrained to move on a mobile or deforming smooth surface (Fig. 5).

In this case, the possible velocity of the particle and, hence, the infinitesimal displacement $d\boldsymbol{r} = \boldsymbol{v}\, dt$ no longer lies in the tangential plane (see Example 2 on page 14). The virtual displacement $\delta \boldsymbol{r}$, which is an infinitesimal possible displacement for a 'stopped' or 'frozen' surface, lies in the tangential plane. Inasmuch as the reaction, in the case of a mobile or deforming smooth surface is directed along the normal to the surface, it follows that $\boldsymbol{R}\,\delta\boldsymbol{r} = 0$ (whereas $\boldsymbol{R}\,d\boldsymbol{r} \neq 0$).

Thus, a smooth surface, whether fixed or mobile or deforming, constitutes an ideal constraint.

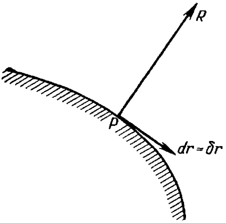

Fig. 4.

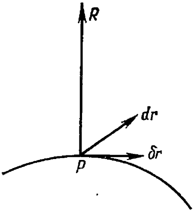

Fig. 5.

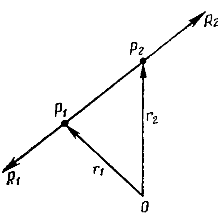

Fig. 6.

Example 2 explains vividly why, when determining nonstationary ideal constraints, it is necessary to equate to zero the work of the reaction forces on arbitrary virtual displacements and not on possible displacements.

In the examples that follow we will encounter only stationary constraints.*

Example 3. Two particles are connected by a rod of invariable length and of negligibly small mass (Fig. 6).

Let us denote by $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ the reaction forces of constraints impressed on particles P_1 and P_2. Then by Newton's third law, the rod is acted upon by the forces $\boldsymbol{N}_1 = -\boldsymbol{R}_1$ and $\boldsymbol{N}_2 = -\boldsymbol{R}_2$. Denoting by m and $\boldsymbol{w}$ the mass of the rod and the acceleration of its centre of inertia, and by I and ε the central moment of inertia and the angular acceleration, we will have

$$m\boldsymbol{w} = \boldsymbol{N}_1 + \boldsymbol{N}_2, \quad I\varepsilon = L$$

where L is the total moment of the forces $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ about the centre of inertia. But it is given that $m = 0$ and $I = 0$. Hence, $\boldsymbol{N}_1 + \boldsymbol{N}_2 = 0$ and $L = 0$.** From these equations it follows that the forces $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ and consequently $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are in direct opposition, i.e. are directed along the rod.

---

* From the definition of ideal constraints it follows that a nonstationary constraint is ideal if all its configurations at different instants (regarded as stationary constraints) are ideal.

** If the motion of the rod is not plane-parallel, the scalar equation $I\varepsilon = L$ is replaced by the vector equation $\frac{d}{dt}(\widehat{\boldsymbol{I}}\omega) = L$, where $\widehat{\boldsymbol{I}}$ is the tensor of inertia and ω is the angular velocity; from $\widehat{\boldsymbol{I}} = 0$ it again follows that $L = 0$.

Further

$$\boldsymbol{R}_1\delta\boldsymbol{r}_1+\boldsymbol{R}_2\delta\boldsymbol{r}_2=\boldsymbol{R}_1\,d\boldsymbol{r}_1+\boldsymbol{R}_2\,d\boldsymbol{r}_2=\boldsymbol{R}_2\,(d\boldsymbol{r}_2-d\boldsymbol{r}_1)=\boldsymbol{R}_2 d\,(\boldsymbol{r}_2-\boldsymbol{r}_1)$$

Let $\boldsymbol{R}_2=c\,(\boldsymbol{r}_2-\boldsymbol{r}_1)$, then

$$\boldsymbol{R}_1\delta\boldsymbol{r}_1+\boldsymbol{R}_2\delta\boldsymbol{r}_2=c\,(\boldsymbol{r}_2-\boldsymbol{r}_1)\,d\,(\boldsymbol{r}_2-\boldsymbol{r}_1)=\frac{c}{2}\,d\,(\boldsymbol{r}_2-\boldsymbol{r}_1)^2=0$$

since $(\boldsymbol{r}_2-\boldsymbol{r}_1)^2=\text{const}$.

It may be taken that an absolutely rigid body is a system of particles in which a constraint of the type considered is imposed on any two particles. For this reason, a rigid body may be regarded as a system of particles subject to ideal constraints. In the absence of other constraints (other than those that effect the rigid connection of the particles of the body) a rigid body is called *free*.

Example 4. Two rigid bodies are connected by a hinge at the point A (Fig. 7). Disregarding the mass and the dimensions of the hinge, it may be asserted (as in the previous example) that $\boldsymbol{R}_1+\boldsymbol{R}_2=0$. Then

$$\boldsymbol{R}_1\delta\boldsymbol{r}+\boldsymbol{R}_2\delta\boldsymbol{r}=(\boldsymbol{R}_1+\boldsymbol{R}_2)\,\delta\boldsymbol{r}=0.$$

Example 5. Two rigid bodies in motion have their ideally smooth surfaces in contact (we disregard friction!) (Fig. 8). In this case too we have $\boldsymbol{R}_1+\boldsymbol{R}_2=0$. Here, $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are directed along the common normal to the surfaces. On the other hand, the relative velocity of these bodies at the point of contact is $\boldsymbol{v}_2-\boldsymbol{v}_1$, and hence the difference of the possible displacements $d\boldsymbol{r}_2-d\boldsymbol{r}_1=(\boldsymbol{v}_2-\boldsymbol{v}_1)\,dt$ lies in the common tangent plane. Therefore

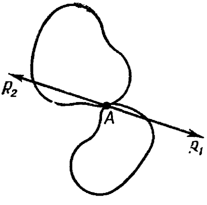

Fig. 7.

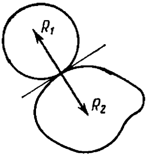

Fig. 8.

$$\boldsymbol{R}_1\delta\boldsymbol{r}_1+\boldsymbol{R}_2\delta\boldsymbol{r}_2=\boldsymbol{R}_1\,d\boldsymbol{r}_1+\boldsymbol{R}_2\,d\boldsymbol{r}_2=\boldsymbol{R}_2\,(d\boldsymbol{r}_2-d\boldsymbol{r}_1)=0$$

Example 6. Two rigid bodies are in motion with their ideally rough (toothed) surfaces in contact. In this case the relative velocity of sliding is $\boldsymbol{v}_2-\boldsymbol{v}_1=0$. Consequently, we also have $d\boldsymbol{r}_2-d\boldsymbol{r}_1=(\boldsymbol{v}_2-\boldsymbol{v}_1)\,dt=0$. Therefore, here too

$$\boldsymbol{R}_1\delta\boldsymbol{r}_1+\boldsymbol{R}_2\delta\boldsymbol{r}_2=\boldsymbol{R}_2\,(d\boldsymbol{r}_2-d\boldsymbol{r}_1)=0$$

A complex mechanism may be regarded as a system of rigid bodies which are connected in pairs either rigidly or by hinges, or the surfaces of which are in contact. If we consider all rigid links absolutely rigid, all articulations as ideal and all contacting planes as ideally smooth or ideally rough, then any complex mechanism may be interpreted as a system of particles that obeys ideal constraints.

It will be noted that in many cases this idealization is not permissible. For example, neglect of the forces of friction may on occasion substantially distort the physical picture of a phenomenon. In such a case, the condition of ideal constraints should be rejected and other conditions should be taken in its place that follow from the nature of the constraints and the laws of friction.

In this interpretation, the concept of ideal constraints becomes almost universally applicable.

From now on it will be assumed that all constraints imposed on a system are ideal.

3. The General Equation of Dynamics. Lagrange's Equations of the First Kind

The following equations hold for particles of a constrained system:

$$m_\nu \boldsymbol{w}_\nu = \boldsymbol{F}_\nu + \boldsymbol{R}_\nu \quad (\nu = 1, \ldots, N) \tag{1}$$

where m_ν is the mass of the νth particle, $\boldsymbol{w}_\nu$ is its acceleration, and $\boldsymbol{F}_\nu$ and $\boldsymbol{R}_\nu$ are, respectively, the resultant of the effective forces and the resultant of the forces of reaction operating on this particle $(\nu = 1, \ldots, N)$. Since the constraints are ideal, for any position of the system under any virtual displacements, we have

$$\sum_{\nu=1}^{N} \boldsymbol{R}_\nu \delta \boldsymbol{r}_\nu = 0. \tag{2}$$

Substituting, here, in place of the reaction forces $\boldsymbol{R}_\nu$ their expressions from equations (1) and multiplying both sides of the resulting equation by -1, we get

$$\sum_{\nu=1}^{N} (\boldsymbol{F}_\nu - m_\nu \boldsymbol{w}_\nu) \delta \boldsymbol{r}_\nu = 0. \tag{3}$$

This is known as the *general equation of dynamics*. It states that, given a system in motion, at any instant of time the *sum of the works of the effective forces and the forces of inertia on any virtual displacements is zero*.

Thus, the general equation of dynamics always holds for any motion that is compatible with constraints and that corresponds to the specified effective forces $\boldsymbol{F}_\nu$ $(\nu = 1, \ldots, N)$.

Conversely, let there now be given a certain motion of a system that is compatible with constraints and for which motion the general equation of dynamics (3) holds. Then, assuming

$$\boldsymbol{R}_\nu = m_\nu \boldsymbol{w}_\nu - \boldsymbol{F}_\nu \quad (\nu = 1, \ldots, N)$$

we will have the equations (1) and (2). Thus, at any instant of time it is possible to choose reaction forces $\boldsymbol{R}_\nu$ such that, by virtue of (2), would be admissible for the given constraints and for which the equations (1) obtained from Newton's second law would hold. We consider that these reaction forces $\boldsymbol{R}_\nu$ are actually realized ("the hypothesis of the realization of admissible reaction forces") and that consequently the motion under consideration corresponds to the given effective forces $\boldsymbol{F}_\nu\ (t, \boldsymbol{r}_\mu, \boldsymbol{v}_\mu)$ $(\nu = 1, \ldots, N)$. Thus, *the general equation of dynamics expresses a necessary and sufficient condition for motion compatible with constraints to correspond to a given system of effective forces* $\boldsymbol{F}_\nu$ $(\nu = 1, \ldots, N)$.*

Let us find expressions for the reaction forces $\boldsymbol{R}_\nu$ by means of the so-called undetermined multipliers of Lagrange. We write out the relations defining the virtual displacements of particles of a system (see Sec. 2):

$$\sum_{\nu=1}^{N} \frac{\partial f_\alpha}{\partial \boldsymbol{r}_\nu} \delta \boldsymbol{r}_\nu = 0 \quad (\alpha = 1, \ldots, d) \tag{4}$$

$$\sum_{\nu=1}^{N} \boldsymbol{l}_{\beta\nu} \delta \boldsymbol{r}_\nu = 0 \quad (\beta = 1, \ldots, g). \tag{5}$$

Multiplying the equations (4) and (5) termwise by arbitrary scalar multipliers $-\lambda_\alpha$ and $-\mu_\beta$ and adding termwise the resulting equations to equation (2), we get

$$\sum_{\nu=1}^{N} \left(\boldsymbol{R}_\nu - \sum_{\alpha=1}^{d} \lambda_\alpha \frac{\partial f_\alpha}{\partial \boldsymbol{r}_\nu} - \sum_{\beta=1}^{g} \mu_\beta \boldsymbol{l}_{\beta\nu} \right) \delta \boldsymbol{r}_\nu = 0. \tag{6}$$

This relation may be written in expanded form as

$$\sum_{\nu=1}^{N} \left(R_{\nu x} - \sum_{\alpha=1}^{d} \lambda_\alpha \frac{\partial f_\alpha}{\partial x_\nu} - \sum_{\beta=1}^{g} \mu_\beta A_{\beta\nu} \right) \delta x_\nu + \{y\}_\nu \, \delta y_\nu + \{z\}_\nu \, \delta z_\nu = 0. \tag{6'}$$

Here, we used $\{y\}_\nu$ and $\{z\}_\nu$ to denote in abbreviated fashion the expressions that differ from the coefficient of δx_ν, given in formula (6′), by replacing the letters x, A with y, B or z and C, respectively.

The relations (7′), Sec. 2, permit expressing $d + g$ out of the $3N$ virtual increments δx_ν, δy_ν, δz_ν in terms of the remaining $n = 3N - d - g$ increments. Here, the determinant J, which is made up of the coefficients of the "dependent" increments in the equations (7′), Sec. 2, is different from zero.

---

* Here, it should be kept in mind that the general equation of dynamics (3) is actually not one equation, but a set of equations because arbitrary virtual displacements may be substituted into equation (3) in place of $\delta \boldsymbol{r}_\nu$ $(\nu = 1, \ldots, N)$ for any instant of time t.

We choose $d+g$ multipliers λ_α and μ_β so that in equation (6′) the coefficients of $d+g$ "independent" increments vanish. This may be done, and done uniquely, since the determinant J of the coefficients of the quantities λ_α, μ_β being determined is not equal to zero. Then the only thing remaining in (6′) are the terms with the independent increments δx_ν, δy_ν, δz_ν. But then the coefficients of these independent increments must also be zero. In other words, the undetermined multipliers λ_α and μ_β may be chosen so that all the scalar coefficients in (6′) and, hence, all the vector coefficients in (6) vanish. But then

$$\boldsymbol{R}_\nu = \sum_{\alpha=1}^{d} \lambda_\alpha \frac{\partial f_\alpha}{\partial \boldsymbol{r}_\nu} + \sum_{\beta=1}^{g} \mu_\beta \boldsymbol{l}_{\beta\nu} \quad (\nu = 1, \ldots, N). \tag{7}$$

We have obtained a general expression for the reaction forces of ideal constraints in terms of the undetermined multipliers of Lagrange λ_α, μ_β ($\alpha = 1, \ldots, d$; $\beta = 1, \ldots, g$).

Putting the expressions (7) for $\boldsymbol{R}_\nu$ into equation (1), we get the so-called *Lagrange equations of the first kind*:*

$$m_\nu \boldsymbol{w}_\nu = \boldsymbol{F}_\nu + \sum_{\alpha=1}^{d} \lambda_\alpha \frac{\partial f_\alpha}{\partial \boldsymbol{r}_\nu} + \sum_{\beta=1}^{g} \mu_\beta \boldsymbol{l}_{\beta\nu} \quad (\nu = 1, \ldots, N) \tag{8}$$

To these equations one should add the constraint equations:

$$f_\alpha(\boldsymbol{r}_\nu) = 0, \quad \sum_{\nu=1}^{N} \boldsymbol{l}_{\beta\nu} \dot{\boldsymbol{r}}_\nu + D_\beta = 0 \quad (\alpha = 1, \ldots, d;\ \beta = 1, \ldots, g). \tag{9}$$

By replacing each vector equation by three scalar equations, we may consider that equations (8) and (9) constitute a set of $3N+d+g$ scalar equations in $3N+d+g$ unknown scalar quantities x_ν, y_ν, z_ν, λ_α, μ_β. Integrating this set we get the final equations of motion and, at the same time, from equations (7) we get the magnitudes of the reaction forces of constraints. However, integration of such a set of equations is ordinarily very involved due to the large number of equations. That is why the Lagrange equations of the first kind find little use in actual practice.

In Secs. 6 and 10 we will derive the Lagrange equations of the second kind for a holonomic system and the Appell equations for a non-holonomic system. In these equations the number of unknown scalar quantities (and, hence, the number of equations) is $3N-d$, i.e. $2d+g$ units less than in the set of equations (8) and (9).

---

* These equations were obtained by the French mathematician and mechanician J. L. Lagrange in his celebrated treatise *Mécanique Analytique* (1788), in which the fundamentals of analytical mechanics were presented for the first time.

Example. Two ponderable particles M_1 and M_2 of identical mass $m=1$ are joined by a rod of invariable length l and negligibly small mass. The system is constrained to move in the vertical plane and only in such manner that the velocity of the midpoint of the rod is directed along it. Determine the motion of the particles M_1 and M_2.

Let x_1, y_1 and x_2, y_2 be the coordinates of the particles M_1 and M_2. We write the constraint equations:

$$\left.\begin{aligned}&\frac{1}{2}\left[(x_2-x_1)^2+(y_2-y_1)^2-l^2\right]=0\\&(x_2-x_1)(\dot y_2+\dot y_1)-(\dot x_2+\dot x_1)(y_2-y_1)=0\end{aligned}\right\}\tag{10}$$

The Lagrange equations with undetermined multipliers λ and μ are of the form

$$\left.\begin{aligned}&\ddot x_1=-\lambda(x_2-x_1)-\mu(y_2-y_1)\\&\ddot y_1=-g-\lambda(y_2-y_1)+\mu(x_2-x_1)\end{aligned}\right\}\tag{11}$$

and

$$\left.\begin{aligned}&\ddot x_2=\lambda(x_2-x_1)-\mu(y_2-y_1)\\&\ddot y_2=-g+\lambda(y_2-y_1)+\mu(x_2-x_1)\end{aligned}\right\}\tag{12}$$

We determine λ and μ from the equations (11) taking into account the first equation of (10):

$$\left.\begin{aligned}&\lambda=-\frac{g}{l^2}(y_2-y_1)-\frac{1}{l^2}\left[(x_2-x_1)\ddot x_1+(y_2-y_1)\ddot y_1\right]\\&\mu=\frac{g}{l^2}(x_2-x_1)-\frac{1}{l^2}\left[(y_2-y_1)\ddot x_1-(x_2-x_1)\ddot y_1\right]\end{aligned}\right\}\tag{13}$$

It will be noted that equations (12) are obtained from (11) if in the latter we replace λ by $-\lambda$ and $\ddot x_1$, $\ddot y_1$ by $\ddot x_2$, $\ddot y_2$. Thus, determining λ and μ from the equations (12), we find

$$\left.\begin{aligned}&\lambda=\frac{g}{l^2}(y_2-y_1)+\frac{1}{l^2}\left[(x_2-x_1)\ddot x_2+(y_2-y_1)\ddot y_2\right]\\&\mu=\frac{g}{l^2}(x_2-x_1)-\frac{1}{l^2}\left[(y_2-y_1)\ddot x_2-(x_2-x_1)\ddot y_2\right]\end{aligned}\right\}\tag{14}$$

Equating the appropriate expressions for μ and λ in the formulae (13) and (14) we get (after simple manipulations)

$$\left.\begin{aligned}&(\ddot x_2-\ddot x_1)(y_2-y_1)-(\ddot y_2-\ddot y_1)(x_2-x_1)=0\\&(\ddot x_2+\ddot x_1)(x_2-x_1)+(\ddot y_2+\ddot y_1)(y_2-y_1)+2g(y_2-y_1)=0\end{aligned}\right\}\tag{15}$$

We introduce the following abbreviated notation:

$$u=x_2-x_1,\quad v=v_2-y_1,\quad P=\dot x_1+\dot x_2,\quad Q=\dot y_1+\dot y_2\tag{16}$$

Then equations (10) and (15) will be written as follows:

$$\left.\begin{aligned} u^2+v^2&=l^2 \\ \ddot{u}v-u\ddot{v}&=0 \end{aligned}\right\} \tag{17}$$

$$Pv-Qu=0 \tag{18}$$

$$\dot{P}u+\dot{Q}v+2gv=0 \tag{19}$$

Equations (17) show that in a u, v-plane a particle with coordinates u, v moves in a circle with radius l and with centre at the origin; its acceleration will all the time be directed towards the centre. The motion of this particle will then be uniform. For this reason,

$$u=l\cos\varphi, \quad v=l\sin\varphi, \quad \dot{\varphi}=\alpha=\text{const}, \quad \varphi=\alpha t+\beta \tag{20}$$

According to (18) we can put

$$P=\frac{f}{l}u, \quad Q=\frac{f}{l}v \tag{21}$$

Substituting these expressions into (19) and taking into account (17) and (20), we find

$$\dot{f}+\frac{2g}{l}v=0, \text{ that is, } \dot{f}=-2g\sin\varphi$$

Then

$$\frac{df}{d\varphi}=\frac{1}{\alpha}\dot{f}=-\frac{2g}{\alpha}\sin\varphi \text{ and } f=\frac{2g}{\alpha}\cos\varphi+2\gamma$$

Consequently, by virtue of (20) and (21), we have

$$P=2\left(\gamma+\frac{g}{\alpha}\cos\varphi\right)\cos\varphi, \quad Q=2\left(\gamma+\frac{g}{\alpha}\cos\varphi\right)\sin\varphi \tag{22}$$

Integrating, we find

$$\left.\begin{aligned} x_1+x_2&=\int P\,dt=\frac{1}{\alpha}\int P\,d\varphi=\frac{2\gamma}{\alpha}\sin\varphi+\frac{g}{\alpha^2}\sin\varphi\cos\varphi+\frac{g}{\alpha^2}\varphi+2\delta \\ y_1+y_2&=-\frac{2\gamma}{\alpha}\cos\varphi-\frac{g}{\alpha^2}\cos^2\varphi+2\varepsilon \end{aligned}\right\} \tag{23}$$

From equations (16), (20) and (23) we finally obtain

$$\left.\begin{aligned} x_1&=\frac{\gamma}{\alpha}\sin\varphi+\frac{g}{2\alpha^2}\sin\varphi\cos\varphi+\frac{g}{2\alpha^2}\varphi-\frac{l}{2}\cos\varphi+\delta \\ y_1&=-\frac{\gamma}{\alpha}\cos\varphi-\frac{g}{2\alpha^2}\cos^2\varphi-\frac{l}{2}\sin\varphi+\varepsilon \\ x_2&=\frac{\gamma}{\alpha}\sin\varphi+\frac{g}{2\alpha^2}\sin\varphi\cos\varphi+\frac{g}{2\alpha^2}\varphi+\frac{l}{2}\cos\varphi+\delta \\ y_2&=-\frac{\gamma}{\alpha}\cos\varphi-\frac{g}{2\alpha^2}\cos^2\varphi+\frac{l}{2}\sin\varphi+\varepsilon \\ \varphi&=\alpha t+\beta \end{aligned}\right\} \tag{24}$$

(where α, β, γ, δ, and ε are arbitrary constants).

4. The Principle of Virtual Displacements. D'Alembert's Principle

An *equilibrium position* is a position of a system such that the system will all the time reside in it if at the initial instant of time it was in that position and the velocities of all its particles were zero.

The position of a system $\boldsymbol{r}_\nu^0$ $(\nu = 1, \ldots, N)$ will be an equilibrium position if and only if the "motion" $\boldsymbol{r}_\nu(t) \equiv \boldsymbol{r}_\nu^0$ $(\nu = 1, \ldots, N)$ satisfies the general equation of dynamics, that is, when

$$\sum_{\nu=1}^{N} \boldsymbol{F}_\nu \delta \boldsymbol{r}_\nu = 0 \tag{1}$$

in this position of the system.

Equation (1) expresses the principle of *virtual displacements. For some position (compatible with constraints) of a system to be an equilibrium position, it is necessary and sufficient that in this position the sum of the works of effective forces on any virtual displacements of the system be zero.*

Ordinarily, the principle of virtual displacements is applied to stationary constraints. If the constraints are stationary, then the term "compatible with constraints" signifies that the position of the system satisfies finite constraints. Now since differential constraints are linear and homogeneous with respect to velocities, they are automatically satisfied since we assume $\boldsymbol{r}_\nu = 0$ $(\nu = 1, \ldots, N)$.

If the constraints are nonstationary, then the term "compatible with constraints" signifies that they are satisfied for any t if in them we put $\boldsymbol{r}_\nu = \boldsymbol{r}_\nu^0$ and $\boldsymbol{v}_\nu = 0$ $(\nu = 1, \ldots, N)$. Note that in this case the virtual displacements $\delta \boldsymbol{r}_\nu$ $(\nu = 1, \ldots, N)$ may also be different for different t.

In the general case, the forces $\boldsymbol{F}_\nu$ depend on t, $\boldsymbol{r}_\mu$, $\boldsymbol{v}_\mu$ $(\mu = 1, \ldots, N)$, that is $\boldsymbol{F}_\nu = \boldsymbol{F}_\nu(t, \boldsymbol{r}_\mu, \boldsymbol{v}_\mu)$ $(\nu = 1, \ldots, N)$. It is then assumed that the equation (1) holds for any value of t if in the expression for $\boldsymbol{F}_\nu$ we put all $\boldsymbol{r}_\mu = \boldsymbol{r}_\mu^0$ and all $\boldsymbol{v}_\mu = 0$.

In the most elementary special cases, the principle of virtual displacements (or, as it is sometimes called when applied to scleronomic systems, the principle of possible displacements) was known at the time of Galileo under the name of the "golden rule of mechanics"*.

Let forces $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ act at the ends of a weightless lever in a state of equilibrium. Then, denoting by F_1' and F_2' the tangential (to possible trajectories) components of these forces, and by δl_1 and δl_2 the magnitudes of the corresponding elementary possible displacements, we will [by virtue of equation (1)] have, except for

---

[*] Galileo attributed substantiation of the "golden rule of mechanics" to Aristotle. In its general statement, the principle of virtual displacements was first given by J. Bernoulli in 1717.

sign,

$$F_1'\delta l_1 = F_2'\delta l_2$$

that is,

$$\frac{\delta l_1}{\delta l_2} = \frac{F_2'}{F_1'}$$

(*a gain in force is compensated for by a loss in displacement and vice versa*—the "golden rule of mechanics").

The principle of virtual displacements is the most general principle of analytical statics. It is possible to obtain from it the conditions of equilibrium of any actual mechanical system.

Example 1. Let us derive from equation (1) the conditions of equilibrium of a free rigid body, which in courses of mechanics are ordinarily obtained through reasoning in geometrical statics. Denoting by $\mathbf{v}_0$ the velocity of some particle of the rigid body, by $\boldsymbol{\omega}$ the angular velocity of the body, by $\mathbf{F}$ and $\mathbf{L}_0$ the principal vector and the principal moment about the pole O for the system of external forces operating on the rigid body, we equate to zero the expression* for the elementary work of forces impressed on the rigid body in an arbitrary infinitesimal displacement of the body:

$$\delta A = (\mathbf{v}_0 + \mathbf{L}_0\boldsymbol{\omega})\, dt\mathbf{F} = 0 \tag{2}$$

Since the vectors $\mathbf{v}_0$ and $\boldsymbol{\omega}$ are arbitrary, equation (2) holds when and only when

$$\mathbf{F} = 0, \text{ and } \mathbf{L}_0 = 0 \tag{3}$$

These equations are the *necessary and sufficient conditions for equilibrium of a free body.*

In similar fashion we obtain the conditions of equilibrium of a constrained rigid body. For example, let point O be fixed. Then $\mathbf{v}_0 = 0$ and equation (2) is of the form $\delta A = \mathbf{L}_0\boldsymbol{\omega}dt = 0$, whence, by virtue of the vector $\boldsymbol{\omega}$ being arbitrary we get the desired equilibrium condition, $\mathbf{L}_0 = 0$.

If the body can only rotate about a fixed axis u (with unit vector $\mathbf{e}$), the equations (2) take on the form $\delta A = \mathbf{L}_0\boldsymbol{\omega}\mathbf{e}\, dt = 0$, whence, by virtue of the arbitrary nature of ω the condition of equilibrium follows: $L_u = 0$; here, $L_u = \mathbf{L}_0\mathbf{e}$ is the principal moment of external forces about the axis u.

Example 2. Let us derive the equilibrium conditions of an arbitrary constrained system of rigid bodies experiencing the action of the force of gravity.

---

* The equation $\delta A = (\mathbf{F}\mathbf{v}_0 + \mathbf{L}_0\boldsymbol{\omega})\, dt$ may be obtained in the following manner. We denote by $\mathbf{F}_i$ the forces acting on particles of the rigid body, by $\mathbf{r}_i$ and $\mathbf{v}_i$ the radii vectors (drawn from point O of the body) and the velocities of the points of application of the forces $\mathbf{F}_i$ $(i = 1, 2, \ldots)$, respectively. Then, using the sign × to indicate vector multiplication, we find

$$\delta A = \sum_i \mathbf{F}_i\delta\mathbf{r}_i = \sum_i \mathbf{F}_i\mathbf{v}_i\, dt = \sum_i \mathbf{F}_i(\mathbf{v}_0 + \boldsymbol{\omega}\times\mathbf{r}_i)\, dt =$$

$$= \Big[\Big(\sum_i \mathbf{F}_i\Big)\mathbf{v}_0 + \boldsymbol{\omega}\sum_i \mathbf{r}_i \times \mathbf{F}_i\Big]\, dt$$

(substitution of $\delta\mathbf{r}_i$ by $d\mathbf{r}_i = \mathbf{v}_i\, dt$ is legitimate due to the fact that the rigid body is a scleronomic system; see page 11). By virtue of Newton's third law, the principal vector and the principal moment of the internal forces in the rigid body are zero. For this reason, $\sum_i \mathbf{F}_i = \mathbf{F}$, and $\sum_i \mathbf{r}_i \times \mathbf{F}_i = \mathbf{L}_0$.

Denote by M the sum of the masses of all bodies and by z_c the vertical coordinate of the centre of gravity of the system of bodies (we take the z-axis to be directed vertically downwards). Then, according to equation (1), we get

$$\delta A = Mg\delta z_c = 0$$

and, consequently, the conditions of equilibrium of the system have the aspect

$$\delta z_c = 0 \tag{4}$$

Thus, the equilibrium positions of a system of rigid bodies are such positions in which the centre of gravity occupies the lowest, the highest or any other "stationary" position in the vertical ("Torricelli's principle").

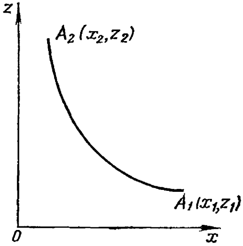

Fig. 9.

Example 3. Find the form of equilibrium of a heavy homogeneous chain attached at two points. Regarding a heavy homogeneous chain as a system of rigid bodies (links) it is possible to write relation (4). But (see Fig. 9, where Oxz is the vertical plane and z is the vertical)

$$z_c = \frac{\int z\,ds}{\int ds}$$

and, since the length of a homogeneous chain does not change during displacements, condition (4) takes the form

$$\delta \int z\,ds = 0 \tag{5}$$

This relation may also be written as

$$\delta \int_{z_1}^{z_2} z \sqrt{1 + \left(\frac{dx}{dz}\right)^2}\,dz = 0 \tag{5'}$$

As is established in the calculus of variations, in the class of curves $x = f(z)$ passing through two specified points, the curve that imparts to the integral

$$\int_{z_1}^{z_2} F\left(z,\ x\ \frac{dx}{dz}\right) dz$$

an extremal (more precisely, stationary) value, for which

$$\delta \int_{z_1}^{z_2} F\, dz = 0,$$

must satisfy the differential equation*

$$\frac{d}{dz}\frac{\partial F}{\partial x'} - \frac{\partial F}{\partial x} = 0 \quad \left(x' = \frac{dx}{dz}\right) \tag{6}$$

In our case $F = z\sqrt{1+\left(\frac{dx}{dz}\right)^2}$. Therefore, equation (6) assumes the form

$$\frac{d}{dz}\left[z\frac{\frac{dx}{dz}}{\sqrt{1+\left(\frac{dx}{dz}\right)^2}}\right] = 0 \tag{7}$$

From this we have

$$z\frac{\frac{dx}{dz}}{\sqrt{1+\left(\frac{dx}{dz}\right)^2}} = c$$

and

$$\frac{dx}{dz} = \frac{c}{\sqrt{z^2 - c^2}} \tag{8}$$

where c is an arbitrary constant. Integrating, we obtain the equation of a catenary:

$$z = \frac{c}{2}\left[e^{(x-\alpha)/c} + e^{-(x-\alpha)/c}\right] = c\cosh\frac{x-\alpha}{c} \tag{9}$$

where the values of the arbitrary constants c and α are determined from the conditions of fixing the ends. Thus, the shape of a homogeneous heavy chain in equilibrium is a catenary.**

Example 4. An invariable plane figure can slide on two of its points A and B along stationary curves lying in the same plane. Let us find out under the action of what force $\boldsymbol{F}$ the figure can be in equilibrium (Fig. 10).

Besides the effective force $\boldsymbol{F}$, the figure is acted upon by two reaction forces directed along normals to the curves, and the lines of action of these three forces have to intersect in a single point. In other words, the line of action of the force $\boldsymbol{F}$ has to pass through the point of intersection of the normals to the curves at the points A and B, i.e. the line of action of the force $\boldsymbol{F}$ has to pass through the instantaneous centre of possible velocities C of the figure.***

---

* This equation was obtained by Euler. For its derivation see page 89 and the note on page 91.

** Galileo thought the parabola to be such a form of equilibrium. Galileo's error was corrected by Huyghens.

*** Here, the magnitude and direction of the force $\boldsymbol{F}$ may be arbitrary.

We can arrive at this conclusion by proceeding from the principle of possible displacements. Indeed, denote by O some point on the line of action of the force $\boldsymbol{F}$. Then from the condition $\delta A = \boldsymbol{F} v_0 \, dt = 0$ we conclude that $v_0 \perp \boldsymbol{F}$, whence it follows that the instantaneous centre of possible velocities of the figure is located on the line of action of the force $\boldsymbol{F}$.

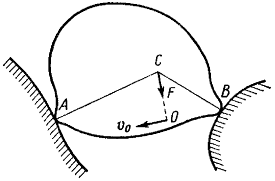

Fig. 10.

Example 5. Some geometrical applications. We begin with a preliminary remark. Let there be given, in a plane, a certain curve C and a point P (the curve C can degenerate into a point in a special case). Draw through P a normal to the curve C and denote by r the distance along the normal from the curve C to the point P. Thus, $r = P_0P$ (Fig. 11). Apply to point P a certain force $\boldsymbol{F}$ directed along the normal P_0P and consider $\boldsymbol{F} > 0$ if the direction of the force $\boldsymbol{F}$ coincides with the direction from P_0 to P, and $F < 0$ otherwise. The elementary work of the force $\boldsymbol{F}$ is equal to $\delta A = \boldsymbol{F}\,(d\boldsymbol{r}_0 + d\boldsymbol{r})$. But $d\boldsymbol{r}$ is composed of two elementary displacements: the displacement along the straight line P_0P (the magnitude of this displacement is dr) and the displacement of the point P caused by rotation of the straight line P_0P. This displacement, like $d\boldsymbol{r}_0$ is perpendicular to the line P_0P, i.e. to the line of action of the force $\boldsymbol{F}$. Therefore,*

$$\delta A = F\,dr \tag{10}$$

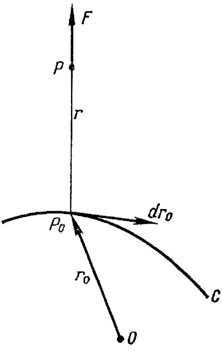

Fig. 11.

Let there be n curves $C_1, C_2, \ldots, C_n$ and a point P in one and the same plane. Denote by $r_1, r_2, \ldots, r_n$ the distances (along the normals) from P to these curves** (Fig. 12 corresponds to the case of $n = 2$). In the same plane we consider the curve D given by the equation***

$$f(r_1, r_2, \ldots, r_n) = 0 \tag{11}$$

---

* When the curve C degenerates into a point, formula (10) yields an expression for the work of a central force.

** Some (or all) of the curves $C_1, C_2, \ldots, C_n$ can degenerate into points.

*** Each of the quantities $r_1, r_2, \ldots, r_n$ is a function of two Cartesian coordinates of the point P. For this reason, equation (11) is the equation of a certain curve in the plane.

We shall now show how, using equation (11), to construct a normal to the curve D at the point P.

For any infinitesimal displacement of point P along the curve D we get

$$\sum_{i=1}^{n} \frac{\partial f}{\partial r_i}\, dr_i = 0 \tag{12}$$

Now apply to point P the forces $F_i = \frac{\partial f}{\partial r_i}$ directed along the normals r_i $(i = 1, \ldots, n)$. Then (12) will be written as

$$\sum_{i=1}^{n} F_i\, dr_i = 0$$

Now this, in accord with the preliminary remark we made, signifies that the sum of the works of the forces $\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_n$ is zero in the case of arbitrary displacement of point P along the curve D. But then any constrained point that can move along the smooth curve D will be in equilibrium under the action of the forces $\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_n$. Therefore, the resultant of the forces $\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_n$ is directed along the normal to the curve D.

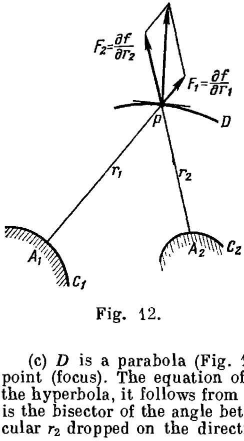
$F_2 = \frac{\partial f}{\partial r_2}$
$F_1 = \frac{\partial f}{\partial r_1}$
P
D
r_1
r_2
A_1
C_1
A_2
C_2

Fig. 12.

We thus have a very simple method for the geometrical construction of a normal to curve D given by equation (11).

Let us consider some special cases:

(a) D is an ellipse. In this case, C_1 and C_2 are points (the foci of the ellipse), equation (11) is of the form $r_1 + r_2 - 2a = 0$, $F_1 = 1$, $F_2 = 1$, and the normal to the ellipse is the bisector of the angle between the focal radii vectors (Fig. 13).

(b) D is a hyperbola. The equation of the hyperbola is $r_1 - r_2 - 2a = 0$, $F_1 = 1$, $F_2 = -1$, and from the construction (Fig. 14) it is easy to see that the tangent to the hyperbola is the bisector of the angle between the focal radii vectors (while the normal is the bisector of the adjacent angle).

(c) D is a parabola (Fig. 15). C_1 is a straight line (the directrix), C_2 is a point (focus). The equation of the parabola is $r_1 - r_2 = 0$. As in the case of the hyperbola, it follows from the construction that the tangent to the parabola is the bisector of the angle between the focal radius vector r_1 and the perpendicular r_2 dropped on the directrix.

For the principle of virtual displacements, equation (1) is a special case of the general equation of dynamics [see equation (3) on page 20]. However, the general equation of dynamics may be regarded as an equation expressing the principle of virtual displacements and characterizing the position of equilibrium of the system that results if to the effective forces $\boldsymbol{F}_\nu$ we add the fictitious inertial forces $-m_\nu \boldsymbol{w}_\nu$ $(\nu = 1, \ldots, N)$. We have thus arrived at *d'Alembert's principle.*

D'Alembert's principle. Any position of a system in motion may be regarded as an equilibrium position if to the effective forces acting on the system in that position we add the fictitious forces of inertia.

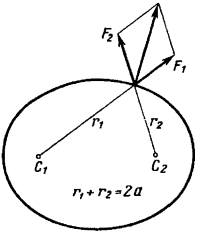

Fig. 13.

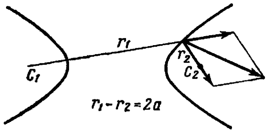

Fig. 14.

D'Alembert's principle permits extending the techniques and methods of solution of static problems to problems of dynamics. In particular, it permits determining dynamic reaction forces by statical methods. Indeed, in an equilibrium position the reaction forces $\boldsymbol{R}_\nu$ differ from $\boldsymbol{F}_\nu - m_\nu \boldsymbol{w}_\nu$ in direction only:

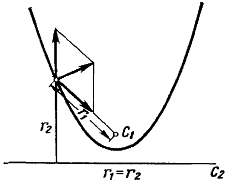

Fig. 15.

$$\boldsymbol{F}_\nu - m_\nu \boldsymbol{w}_\nu = -\boldsymbol{R}_\nu \quad (\nu = 1, \ldots, N).$$

But then

$$m_\nu \boldsymbol{w}_\nu = \boldsymbol{F}_\nu + \boldsymbol{R}_\nu \quad (\nu = 1, \ldots, N),$$

i.e. the reaction forces $\boldsymbol{R}_\nu$ determined by means of the d'Alembert principle are the desired dynamical reaction forces. And so we can add the following proposition to the above formulated d'Alembert principle:

When regarding inertial forces as additional effective forces applied to the points of a system, we replace the given dynamical problem by a new statical problem. In the new problem, the static reaction forces coincide with the sought-for reaction forces in the initial dynamical problem.

The following examples will serve to illustrate the use of statical methods in the solution of problems in dynamics.

Example 1. A locomotive tender filled with water is in motion with an acceleration w. It is required to determine the shape and position of the water surface.

In the absence of acceleration the surface of the water is horizontal. The given plane is perpendicular, at each point, to the direction of the body forces

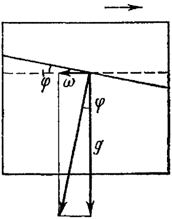

Fig. 16.

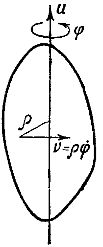

Fig. 17.

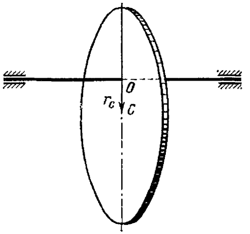

Fig. 18.

of gravity applied to the water. This static position is also applicable to the case of the accelerated motion of the tender if to each element of mass dm there is applied an additional fictitious force of inertia $dJ = -dmw$. The surface of water will be a plane perpendicular to the resultant of the two body forces: the vertical force of the gravity dmg and the horizontal force of inertia $-dmw$ (Fig. 16). The water surface will be inclined to the horizon at an angle φ, where $\tan\varphi = \frac{w}{g}$.

Example 2. Let us write the differential equation of rotation of a rigid body about a fixed axis u (Fig. 17). To each element of mass dm we apply a fictitious inertial force $-dmw$. Calculate the principal moment of inertia forces about the axis of rotation

$$-\int dm\rho\ddot{\varphi}\rho = -\ddot{\varphi}\int \rho^2\, dm = -I_u\ddot{\varphi}$$

where $I_u = \int \rho^2\, dm$ is the moment of inertia of the body about the rotation axis u. Denote by L_u the principal moment of the external forces applied to the body about the axis u.* Then, according to d'Alembert's principle the body will be in equilibrium under the action of the total moment $L_u - I_u\ddot{\varphi}$. Therefore, (see page 26) this total moment must equal zero. We get

$$I_u\ddot{\varphi} = L_u$$

Example 3. A horizontal homogeneous shaft is in uniform rotation with angular velocity ω. A homogeneous disc is mounted eccentrically on the shaft perpendicular to the axis of the shaft at equal distances from the bearings. It is required to determine the pressure on the bearings during rotation of the shaft.

---

* The principal moment of the internal forces is zero.

We consider the inertial forces $dm\omega^2 \boldsymbol{r}$ that correspond to the separate elements dm of the disc (Fig. 18). These are converging forces directed away from the axis of the shaft. The resultant of these forces is equal to $\boldsymbol{J} = \omega^2 \int \boldsymbol{r}\, dm = M_1\omega^2 \boldsymbol{r}_C$, where M_1 is the mass of the disc, and $\boldsymbol{r}_C = \boldsymbol{OC}$ (O is the point of intersection of the plane of the disc with the axis of the shaft, and C is the geometric centre of the disc). Apply the d'Alembert principle and determine the static pressure on the bearings, assuming that three forces are applied to the axis of the shaft:* (1) the force of gravity of the shaft $M\boldsymbol{g}$; (2) the force of gravity of the disc $M_1\boldsymbol{g}$, and (3) the force $\boldsymbol{J} = M_1\omega^2\boldsymbol{r}_C$.

The pressure $\boldsymbol{N}$ on each bearing is determined from the formula

$$\boldsymbol{N} = \frac{1}{2}(m + M_1)\,\boldsymbol{g} + \frac{1}{2} M_1\omega^2 \boldsymbol{r}_C$$

The force $\boldsymbol{N}$ has a maximum value

$$N_{max} = \frac{1}{2}(m + M_1)\, g + \frac{1}{2} M_1\omega^2 OC$$

in the position of the disc when the geometric centre of the disc C is located under the point O.

5. Holonomic Systems. Independent Coordinates. Generalized Forces

Let there be given a holonomic system of N particles P_ν with radii vectors $\boldsymbol{r}_\nu = x_\nu \boldsymbol{i} + y_\nu \boldsymbol{j} + z_\nu \boldsymbol{k}$ ($\nu = 1, \ldots, N$)** subject to the finite constraints

$$f_\alpha(t, \boldsymbol{r}_\nu) = 0 \quad (\alpha = 1, \ldots, d) \tag{1}$$

or (in equivalent notation)

$$f_\alpha(t, x_\nu, y_\nu, z_\nu) = 0 \quad (\alpha = 1, \ldots, d) \tag{1'}$$

We shall assume that d functions f_α of $3N$ arguments x_ν, y_ν, z_ν ($\nu = 1, \ldots, N$) are independent ***; here t is regarded as a parameter. We can therefore express d coordinates of equations (1′) as the functions of $3N - d$ of the remaining ones and the time t and regard these $3N - d$ coordinates as independent quantities that define the position of the system at time t.

---

* The resultant of the elementary forces of inertia for a shaft is zero and is therefore ignored.

** $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ are the unit vectors of the x-, y-, and z-axes of the inertial system of coordinates.

*** Otherwise, if we have, for example, a relation like

$$f_d = \Omega(f_1, \ldots, f_{d-1}, t),$$

one of the constraints (in the given case, $f_d = 0$) would either contradict the others [for $\Omega(0, \ldots, 0, t) \not\equiv 0$], or would be a consequence of the others [for $\Omega(0, \ldots, 0, t) \equiv 0$].

However, one need not take Cartesian coordinates for such independent coordinates. All the $3N$ Cartesian coordinates may be expressed in the form of functions of $n = 3N - d$ independent parameters $q_1, \ldots, q_n$ and of t:

$$\left.\begin{aligned} x_\nu = \varphi_\nu(t, q_1, \ldots, q_n),\ y_\nu = \psi_\nu(t, q_1, \ldots, q_n), \\ z_\nu = \chi_\nu(t, q_1, \ldots, q_n) \quad (\nu = 1, \ldots, N). \end{aligned}\right\} \tag{2}$$

When these functions are put in the constraint equations (1′), the latter become identities. We will assume that any position of the system that is compatible with constraints at the given instant of time may be obtained from the equations (2) for certain values of the quantities $q_1, \ldots, q_n$.

The equations (2) are equivalent to the vector equations

$$\boldsymbol{r}_\nu = \boldsymbol{r}_\nu(t, q_1, \ldots, q_n) \quad (\nu = 1, \ldots, N) \tag{2′}$$

The scalar functions (2) and consequently the vector functions (2′) are assumed continuous and differentiable.

The minimal number of quantities q_i with the aid of which formulas (2) can embrace all possible positions of a holonomic system coincides with the number of degrees of freedom of the system $n = 3N - d$ (see page 15).

The quantities $q_1, \ldots, q_n$ in formulas (2) or (2′) (n is the number of degrees of freedom) are called the independent generalized coordinates of the system.

For each instant of time t, a one-to-one correspondence is established between the possible states of the system and the points of a certain region in the n-dimensional coordinate space $(q_1, \ldots, q_n)$. To each position of the system at time t there corresponds a point in the space $(q_1, \ldots, q_n)$ that describes this position of the system. The motion of a point in the coordinate space $(q_1, \ldots, q_n)$ corresponds to the motion of the system.

If all constraints are stationary (a scleronomic system!), then the time t does not appear explicitly in equations (1′). It is then always possible to choose coordinates $q_1, \ldots, q_n$ such that the time t does not enter the equations (2) either. From now on it is assumed that for a scleronomic system the independent coordinates $q_1, \ldots, q_n$ are chosen in precisely that way. Then, for a scleronomic system, the formulas (2) and (2′) take on the form

$$x_\nu = \varphi_\nu(q_i),\ y_\nu = \psi_\nu(q_i),\ z_\nu = \chi_\nu(q_i) \quad (\nu = 1, \ldots, N) \tag{3}$$

or

$$\boldsymbol{r}_\nu = \boldsymbol{r}_\nu(q_i) \quad (\nu = 1, \ldots, N) \tag{3′}$$

Example 1. A double pendulum (Fig. 19) moving in a plane has two degrees of freedom. For the independent coordinates q_1 and q_2 we can take the angles φ and ψ.

Example 2. A free rigid body has six degrees of freedom. For the independent coordinates we can take the three coordinates x_A, y_A, z_A of some point A of the body and the three Eulerian angles ψ, θ and φ that define the rotation of a system of axes $A\xi\eta\zeta$ (that is invariably fixed to the body) about a stationary system of coordinate axes $Oxyz$.

Euler's angles are determined as follows (Fig. 20). Draw through point A the axes Ax_1, Ay_1, Az_1 parallel with the axes Ox, Oy, Oz and in the same directions. The line AN that intersects the planes Ax_1y_1 and $A\xi\eta$ is called the nodal line.* Then θ, or the angle of nutation, is the angle between the axes Az_1 and $A\zeta$; ψ, or the angle of precession, is the angle between the axes Ax_1 and AN; and φ is the angle of pure rotation formed by the axes AN and $A\xi$.

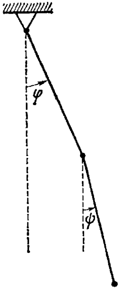

Fig. 19.

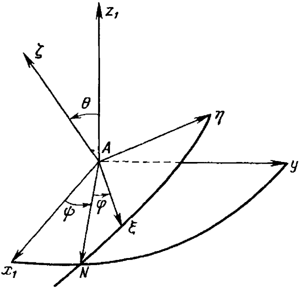

Fig. 20.

By means of three parallel displacements—along the x-, y- and z-axis—by x_A, y_A, z_A the coordinate trihedral $Oxyz$ is displaced to the position $Ax_1y_1z_1$. By means of three successive rotations—through the angle ψ about the axis Az_1, through the angle θ about the axis AN, and through the angle φ about the axis $A\zeta$—the trihedral $Ax_1y_1z_1$ is displaced to the position $A\xi\eta\zeta$.

Thus, the quantities x_A, y_A, z_A, ψ, θ, φ determine the position of the coordinate trihedral $A\xi\eta\zeta$ relative to the trihedral $Oxyz$, i.e. they determine the position of the given rigid body relative to the initial system of coordinate axes.

Take an arbitrary point of a rigid body. It is determined by specifying its coordinates ξ, η, ζ. Then the coordinates x, y, z of this point may be represented as functions of the quantities x_A, y_A, z_A, ψ, θ, φ. Thus, for example, from Fig. 20 it will readily be seen that $z = z_A + \xi \sin\varphi \sin\theta + \eta \cos\varphi \sin\theta + \zeta \cos\theta$.

Similar, though somewhat more involved formulas are obtained for x and y [24]. These formulas are a special case of the formulas (2). They do not contain t explicitly. *A free rigid body is a scleronomic system.*

It will be noted that in the motion of a rigid body the quantities x_A, y_A, z_A, ψ, θ, φ vary, and the decomposition given above of the transition from

---

* Direct the axis AN so that rotation about this axis from the axis Az_1 to $A\zeta$ through the smallest angle is counterclockwise.

$Oxyz$ to $A\xi\eta\zeta$ into three parallel displacements and three rotations gives the impression of arbitrary motion of the rigid body in the form of complex (compound) motion consisting of six simple motions: three translational motions (along the x-, y-, and z-axis) and three purely rotational motions (about the axes Az_1, AN, and $A\zeta$). Since in complex motion the angular velocity is equal to the vector sum of the component angular velocities, it follows that

$$\boldsymbol{\omega} = \boldsymbol{\omega}_{\psi} + \boldsymbol{\omega}_{\theta} + \boldsymbol{\omega}_{\varphi} \tag{4}$$

where $\boldsymbol{\omega}_{\psi}$, $\boldsymbol{\omega}_{\theta}$, $\boldsymbol{\omega}_{\varphi}$ are directed along the axes Az_1, AN, $A\zeta$, respectively; here, $\omega_{\psi} = \dot{\psi}$, $\omega_{\theta} = \dot{\theta}$; and $\omega_{\varphi} = \dot{\varphi}$.

Example 3. A free particle M has three degrees of freedom. For the independent coordinates we can take Cartesian coordinates or some other coordinates of the point. For the case when cylindrical coordinates r, ψ, z are taken for q_1, q_2, q_3, the formulas (2) will appear as (Fig. 21):

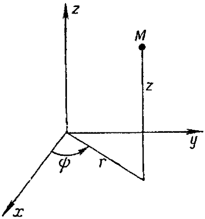

Fig. 21.

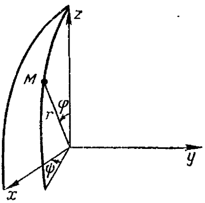

Fig. 22.

$$x = r\cos\psi, \quad y = r\sin\psi, \quad z = z \tag{5}$$

In the case of the spherical coordinates r, φ, ψ (Fig. 22), we will have, in place of formulas (5),

$$x = r\cos\psi\sin\varphi, \quad y = r\sin\psi\sin\varphi, \quad z = r\cos\varphi \tag{6}$$

Example 4. A constrained particle M is located on a mobile sphere

$$(x = at)^2 + (y - bt)^2 + (z - ct)^2 = r^2$$

Then $n = 2$, and the longitude and latitude on the sphere (Fig. 23) may be utilized as the independent coordinates:

$$x = at + r\cos q_1 \cos q_2, \quad y = bt + r\sin q_1 \cos q_2,$$
$$z = ct + r\sin q_2$$

To every coordinate q_i there corresponds a generalized force Q_i ($i = 1, \ldots, n$). The generalized forces are determined as follows. Consider the elementary work of effective forces on virtual displacements

$$\delta A = \sum_{\nu=1}^{N} \boldsymbol{F}_{\nu}\delta\boldsymbol{r}_{\nu} \tag{7}$$

But the virtual differentials, i.e. the differentials for fixed "frozen" t, of the function $\boldsymbol{r}_\nu(t, q_i)$* are the virtual displacements $\delta\boldsymbol{r}_\nu$:

$$\delta\boldsymbol{r}_\nu = \sum_{i=1}^{n} \frac{\partial \boldsymbol{r}_\nu}{\partial q_i}\delta q_i \quad (\nu = 1, \ldots, N) \tag{8}$$

Substitute the expressions (8) into the right-hand side of formula (7) and express the elementary work of the effective forces

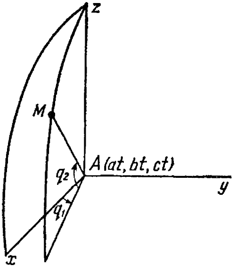

Fig. 23.

on the virtual displacements in terms of arbitrary elementary increments δq_i of the independent coordinates q_i $(i = 1, \ldots, n)$:

$$\delta A = \sum_{\nu=1}^{N} \boldsymbol{F}_\nu \sum_{i=1}^{n} \frac{\partial \boldsymbol{r}_\nu}{\partial q_i}\delta q_i = \sum_{i=1}^{n}\left(\sum_{\nu=1}^{N} \boldsymbol{F}_\nu \frac{\partial \boldsymbol{r}_\nu}{\partial q_i}\right)\delta q_i = \sum_{i=1}^{n} Q_i\delta q_i \tag{9}$$

where the coefficients of δq_i, or the generalized forces Q_i, are determined by the equations

$$Q_i = \sum_{\nu=1}^{N} \boldsymbol{F}_\nu \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \quad (i = 1, \ldots, n) \tag{10}$$

---

* Indeed, when the functions $r_\nu(t, q_i)$ $(\nu = 1, \ldots, N)$ are substituted into the constraint equations $f_\alpha(t, \boldsymbol{r}_\nu) = 0$ $(\alpha = 1, \ldots, d)$, the latter become identities. Let us differentiate termwise the identities obtained (for fixed t). We then find

$$\sum_{\nu=1}^{N} \frac{\partial f_\alpha}{\partial r_\nu}\delta\boldsymbol{r}_\nu = 0 \ (\alpha = 1, \ldots, d) \tag{*}$$

where δr_ν $(\nu = 1, \ldots, N)$ are the virtual differentials. But the equations (*) coincide with the first d equations (7) on page 13 that determine the virtual displacements of a holonomic system. Consequently, the virtual differentials of the radii vectors are virtual displacements of points of a holonomic system.

It will be noted that for practical purposes formula (10) is by far not always used to find the quantity Q_i. Instead, the system is given an elementary virtual displacement such that only the ith coordinate q_i receives a certain increment while the remaining independent coordinates do not change. After that the work of the effective forces δA_i is calculated on just such a specially chosen displacement. Then,

$$\delta A_i = Q_i \delta q_i, \quad \text{and} \quad Q_i = \frac{\delta A_i}{\delta q_i}$$

Example 5. A rigid body is constrained to move translationally along the x-axis. Then $n = 1$ and the abscissa x of some point of the body A may be taken as the independent coordinate. Here,

$$\delta A = X \delta x \tag{11}$$

where X is the sum of the projections, on the x-axis, of all effective forces acting on the body. Obviously X is the generalized force for the coordinate x:

$$Q = X \tag{12}$$

Example 6. A rigid body is constrained to rotate about a certain fixed axis u. The angle of rotation φ may be taken as the independent coordinate. Then

$$\delta A = L_u \delta \varphi \tag{13}$$

where L_u is the total moment of all effective forces about the axis of rotation, and

$$Q = L_u \tag{14}$$

Example 7. A free rigid body. For the independent coordinates let us take the three coordinates x_A, y_A, z_A of some point A of the body and the three Eulerian angles ψ, θ, φ (see Example 2 on pages 35-36). Then, according to equation (9),

$$\delta A = Q_x \delta x + Q_y \delta y + Q_z \delta z + Q_\psi \delta \psi + Q_\theta \delta \theta + Q_\varphi \delta \varphi \tag{15}$$

To determine Q_x we impart to the body an elementary displacement along the x-axis. Then $\delta y_A = \delta z_A = 0$, and $\delta \psi = \delta \theta = \delta \varphi = 0$. Therefore, $\delta A = Q_x \delta x_A$. A comparison with (11) yields

$$Q_x = X$$

Similarly, $Q_y = Y$, $Q_z = Z$. Here, X, Y, Z are the projections on the stationary axes x, y, z of the principal vector of all effective forces acting on the body.

We now impart to our body an elementary displacement such that only the angle ψ changes, while the quantities x_A, y_A, z_A, θ and φ remain invariable. Then

$$\delta A = Q_\psi \delta \psi$$

On the other hand, the elementary displacement of the body is a rotation about the axis Az_1. Therefore, in accord with formula (13),

$$Q_\psi = L_\psi$$

where L_ψ is the total moment of all effective forces about the axis Az_1, about which a rotation through the angle ψ is performed.

In quite analogous fashion, $Q_\theta = L_\theta$ and $Q_\varphi = L_\varphi$, where L_θ and L_φ are the total moments of the effective forces about the axes AN and $A\zeta$.

We can arrive at the same expressions for the generalized forces if we take advantage of the expression for the elementary work of the effective forces applied to a rigid body (see page 26):*

$$\delta A = \boldsymbol{R}\delta\boldsymbol{r}_A + \boldsymbol{L}_A\boldsymbol{\omega}\,dt \tag{16}$$

where $\boldsymbol{R}$ and $\boldsymbol{L}_A$ are the principal vector and the principal moment of the system of forces about the pole A. Inasmuch as [see formula (4)] $\boldsymbol{\omega} = \boldsymbol{\omega}_\psi + \boldsymbol{\omega}_\theta + \boldsymbol{\omega}_\varphi$, where $\omega_\psi = \dot{\psi}$, $\omega_\theta = \dot{\theta}$, $\omega_\varphi = \dot{\varphi}$, and the projections of the vector $\boldsymbol{L}_A$ on the directions of the vectors $\boldsymbol{\omega}_\psi$, $\boldsymbol{\omega}_\theta$, $\boldsymbol{\omega}_\varphi$ are equal, respectively, to L_ψ, L_θ, L_φ, from formula (16) we find

$$\delta A = X\delta x_A + Y\delta y_A + Z\delta z_A + L_\psi\delta\psi + L_\theta\delta\theta + L_\varphi\delta\varphi \tag{17}$$

A comparison of the expressions (17) and (15) gives us expressions for the generalized forces.

Now let a certain position of the system be a position of equilibrium. According to the principle of virtual displacements this is possible if and only if

$$\delta A = \sum_{i=1}^{n} Q_i\delta q_i = 0 \tag{18}$$

But the increments δq_i in the independent coordinates q_i may be quite arbitrary. Therefore the equation (18) is equivalent to the system of equations

$$Q_i = 0 \quad (i = 1, \ldots, n). \tag{19}$$

Thus, *the position of a holonomic system is an equilibrium position if and only if all the generalized forces in this position are zero.*

Example 8. In accordance with the equations (19), the conditions of equilibrium of a free rigid body may be written as

$$X = Y = Z = 0, \quad L_\psi = L_\theta = L_\varphi = 0 \tag{20}$$

(see preceding example). Here, X, Y, Z are the projections, on the coordinate axes, of the principal vector $\boldsymbol{R}$ of the external forces acting on the body, and L_ψ, L_θ, L_φ are the projections of the principal moment $\boldsymbol{L}_A$ of these forces on three non-coplanar directions. For this reason, the scalar equations (20) are equivalent to two vector equations:

$$\boldsymbol{R} = 0, \quad \boldsymbol{L}_A = 0$$

These are the necessary and sufficient conditions of equilibrium of a free rigid body that had already been established on page 26.

---

* Since we have to do here with a scleronomic system, in place of δ we can write d, and vice versa. Therefore, $d\boldsymbol{r}_A = \delta\boldsymbol{r}_A$ and $\delta\psi = d\psi = \dot{\psi}\,dt$, $\delta\theta = \dot{\theta}\,dt$ and $\delta\varphi = \dot{\varphi}\,dt$.

6. Lagrange's Equations of the Second Kind in Independent Coordinates

In deriving the differential equations of motion of a holonomic system in independent coordinates $q_1, \ldots, q_n$, we shall proceed from the general equation of dynamics

$$\sum_{\nu=1}^{N} (\boldsymbol{F}_\nu - m_\nu \boldsymbol{w}_\nu)\, \delta \boldsymbol{r}_\nu = 0 \tag{1}$$

Let us recall the expression, obtained in Sec. 5, for the elementary work of effective forces:

$$\delta A = \sum_{\nu=1}^{N} \boldsymbol{F}_\nu \delta \boldsymbol{r}_\nu = \sum_{i=1}^{n} Q_i \delta q_i \tag{2}$$

where

$$Q_i = \sum_{\nu=1}^{N} \boldsymbol{F}_\nu \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \quad (i = 1, \ldots, n) \tag{3}$$

Quite analogously it is possible to represent the elementary work of the inertial forces $-m_\nu \boldsymbol{w}_\nu$ $(\nu = 1, \ldots, N)$:

$$\delta A_J = -\sum_{\nu=1}^{N} m_\nu \boldsymbol{w}_\nu \delta \boldsymbol{r}_\nu = -\sum_{i=1}^{n} Z_i \delta q_i \tag{4}$$

where, by analogy with expression (3),

$$Z_i = \sum_{\nu=1}^{N} m_\nu \boldsymbol{w}_\nu \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} = \sum_{\nu=1}^{N} m_\nu \frac{d\dot{\boldsymbol{r}}_\nu}{dt} \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} =$$

$$= \frac{d}{dt} \sum_{\nu=1}^{N} m_\nu \dot{\boldsymbol{r}}_\nu \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} - \sum_{\nu=1}^{N} m_\nu \dot{\boldsymbol{r}}_\nu \frac{d}{dt} \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \quad (i = 1, \ldots, n) \tag{5}$$

But the velocity

$$\dot{\boldsymbol{r}}_\nu = \sum_{k=1}^{n} \frac{\partial \boldsymbol{r}_\nu}{\partial q_k} \dot{q}_k + \frac{\partial \boldsymbol{r}_\nu}{\partial t} \tag{6}$$

is linearly dependent on $\dot{q}_k$ $(k = 1, \ldots, n)$. From this formula we find that

$$\frac{\partial \dot{\boldsymbol{r}}_\nu}{\partial \dot{q}_i} = \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \quad (i = 1, \ldots, n;\ \nu = 1, \ldots, N) \tag{7}$$

On the other hand, from the same equation (6) we obtain

$$\frac{\partial \dot{\boldsymbol{r}}_\nu}{\partial q_i} = \sum_{k=1}^{n} \frac{\partial^2 \boldsymbol{r}_\nu}{\partial q_i\, \partial q_k} \dot{q}_k + \frac{\partial^2 \boldsymbol{r}_\nu}{\partial q_i\, \partial t} = \frac{d}{dt} \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \tag{8}$$

$$(i = 1, \ldots, n;\ \nu = 1, \ldots, N)$$

Therefore, expression (5) for Z_i may also be written as follows:

$$Z_i = \frac{d}{dt} \sum_{\nu=1}^{N} m_\nu \dot{\boldsymbol{r}}_\nu \frac{\partial \dot{\boldsymbol{r}}_\nu}{\partial \dot{q}_i} - \sum_{\nu=1}^{N} m_\nu \dot{\boldsymbol{r}}_\nu \frac{\partial \dot{\boldsymbol{r}}_\nu}{\partial q_i} = \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i}$$

$$(i = 1, \ldots, n) \tag{9}$$

where T is the kinetic energy of the system:

$$T = \frac{1}{2} \sum_{\nu=1}^{N} m_\nu \dot{\boldsymbol{r}}_\nu^2 \tag{10}$$

The general equation of dynamics (1) gives us

$$\delta A + \delta A_J = 0 \tag{11}$$

or, by virtue of the equations (2) and (4),

$$\sum_{i=1}^{n} (Q_i - Z_i)\, \delta q_i = 0 \quad (i = 1, \ldots, n) \tag{12}$$

Since the q_i are independent coordinates and, for this reason, the δq_i are absolutely arbitrary increments in the coordinates ($i = 1, \ldots, n$), it follows that (12) can hold when and only when all the coefficients of δq_i in equation (12) are equal to zero. Therefore, the general equation of dynamics (12) is equivalent to the set of equations

$$Z_i = Q_i \quad (i = 1, \ldots, n) \tag{13}$$

which, according to the relations (9), may also be written in the following form:

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = Q_i \quad (i = 1, \ldots, n)\,. \tag{14}$$

Equations (14) are called the *Lagrange equations of the second kind* or the *Lagrange equations in independent coordinates.*

The quantities $\dot{q}_i$ ($i = 1, \ldots, n$) are called *generalized velocities.* The velocities of points of the system $\boldsymbol{v}_\nu = \dot{\boldsymbol{r}}_\nu$ are expressed in terms of the generalized velocities (and also in terms of independent coordinates and the time) by means of the formulas (6). The quantities $\ddot{q}_i$ ($i = 1, \ldots, n$) are called *generalized accelerations.*

After performing the operation $\frac{d}{dt}$, the left-hand sides of the Lagrange equations (14) contain the time t, the generalized coordinates q_i, the generalized velocities $\dot{q}_i$, and the generalized accelerations $\ddot{q}_i$ $(i = 1, \ldots, n)$. The generalized forces Q_i $(i = 1, \ldots, n)$ on the right-hand sides of the Lagrange equations are ordinarily specified * as functions of t, q_k, $\dot{q}_k$ $(k = 1, \ldots, n)$:

$$Q_i = Q_i (t, q_k, \dot{q}_k) \quad (i = 1, \ldots, n) \tag{15}$$

The Lagrange equations (14) form a set of n ordinary differential equations of the second order in n unknown functions q_i of the independent variable t. The order of this system is $2n$. Note that the set of differential equations determining the motion of a holonomic system with n degrees of freedom cannot be of order less than $2n$, since by virtue of the arbitrariness of the initial values of the quantities q_i and $\dot{q}_i$ $(i = 1, \ldots, n)$, the solution of the system must contain at least $2n$ arbitrary constants. Thus, the set of Lagrange equations in independent coordinates has the lowest possible order.

In the case of a constrained system, the reaction forces $\boldsymbol{R}_\nu$ $(\nu = 1, \ldots, N)$ have likewise to be determined. The reaction forces do not enter into Lagrange's equations. This is an essential advantage of Lagrange's equations. After Lagrange's equations have been integrated and the functions $q_i(t)$ $(i = 1, \ldots, n)$ found, $\boldsymbol{r}_\nu = \boldsymbol{r}_\nu(t)$ are determined [by substitution of these functions into formulas (2′) on page 34] and, consequently, $\boldsymbol{v}_\nu = \dot{\boldsymbol{r}}_\nu$, $\boldsymbol{w}_\nu = \ddot{\boldsymbol{r}}_\nu$ and $\boldsymbol{F}_\nu (t, \boldsymbol{r}_\nu, \dot{\boldsymbol{r}}_\nu)$ $(\nu = 1, \ldots, N)$. After that the unknown reaction forces are determined from the formulas

$$\boldsymbol{R}_\nu = m_\nu \boldsymbol{w}_\nu - \boldsymbol{F}_\nu \quad (\nu = 1, \ldots, N) \tag{16}$$

In the case of a free system of particles, Lagrange's equations are a compact notation of the equations of motion in an arbitrary system of coordinates.

Example 1. A rigid body is in rotation about a stationary axis u. For the independent coordinate we take the angle of rotation φ. The appropriate generalized force Q (see Example 6 on page 38) is equal to the moment of rotation L_u. On the other hand, $T = \frac{1}{2} I_u \dot{\varphi}^2$ where I_u is the moment of inertia of the body about the axis of rotation.

---

* See formulas (3) and (6) of this section and also formula (10) on page 16 and formula (2′) on page 34.

The Lagrange equation

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{\varphi}}-\frac{\partial T}{\partial \varphi}=Q$$

after the substitution

$$\frac{\partial T}{\partial \dot{\varphi}}=I_u\dot{\varphi},\quad \frac{\partial T}{\partial \varphi}=0,\quad Q=L_u$$

assumes the form

$$I_u\ddot{\varphi}=L_u$$

This is the differential equation of the rotation of a rigid body about a stationary axis.

Example 2. A double simple pendulum is in motion in a plane (Fig. 24). We form an expression for the elementary work:

$$\delta A=m_1g\delta z_1+m_2g\delta z_2$$

where $z_1=l_1\cos\varphi_1$, $z_2=l_1\cos\varphi_1+l_2\cos\varphi_2$.

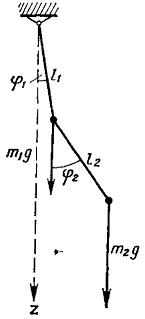

Fig. 24.

Calculating δz_1 and δz_2, we find

$$\delta A=-(m_1+m_2)\,gl_1\sin\varphi_1\delta\varphi_1-m_2gl_2\sin\varphi_2\delta\varphi_2$$

and

$$Q_1=-(m_1+m_2)\,gl_1\sin\varphi_1,\quad Q_2=-m_2gl_2\sin\varphi_2$$

On the other hand,

$$T=\frac{1}{2}m_1l_1^2\dot{\varphi}_1^2+\frac{1}{2}m_2\,[l_1^2\dot{\varphi}_1^2+l_2^2\dot{\varphi}_2^2+2l_1l_2\cos(\varphi_1-\varphi_2)\,\dot{\varphi}_1\dot{\varphi}_2]=$$

$$=\frac{1}{2}(m_1+m_2)\,l_1^2\dot{\varphi}_1^2+m_2l_1l_2\dot{\varphi}_1\dot{\varphi}_2\cos(\varphi_1-\varphi_2)+\frac{1}{2}m_2l_2^2\dot{\varphi}_2^2$$

The first Lagrange equation

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{\varphi}_1}-\frac{\partial T}{\partial \varphi_1}=Q_1$$

is of the form

$$\frac{d}{dt}\left[(m_1+m_2)\, l_1^2\dot{\varphi}_1+m_2l_1l_2\dot{\varphi}_2\cos(\varphi_1-\varphi_2)\right]+m_2l_1l_2\dot{\varphi}_1\dot{\varphi}_2\sin(\varphi_1-\varphi_2)=$$

$$=-(m_1+m_2)\,gl_1\sin\varphi_1$$

It is left to the reader to form the second equation that corresponds to the coordinate φ_2.

Example 3. It is required to determine the differential equations of motion of a free particle in spherical coordinates (see Example 3 on page 36, and Fig. 22). The velocity of the particle is equal to the vector sum of: (a) radial velocity; (b) rotatory velocity due to the rotation of the radius in the plane of the meridian, and (c) rotatory velocity due to rotation of the plane of meridian. The velocity components are pairwise orthogonal and therefore

$$T=\frac{1}{2}\,mv^2=\frac{1}{2}\,m\,(\dot{r}^2+r^2\dot{\varphi}^2+r^2\sin^2\varphi\dot{\psi}^2)$$

To find the generalized force Q_r, we displace the particle along the radius. Then $\delta A_r = F_r\delta r$, where F_r is the projection of the applied force $\boldsymbol{F}$ on the direction of the radius. Whence $Q_r = F_r$.

Now let us impart to the particle an elementary displacement along the meridian. Then $\delta A_\varphi = F_\varphi r\delta\varphi$, where F_φ is the projection of the force $\boldsymbol{F}$ on the tangent to the meridian.* Therefore

$$Q_\varphi=F_\varphi r$$

Analogously,

$$Q_\psi=F_\psi r\sin\varphi$$

where F_ψ is the projection of the force $\boldsymbol{F}$ on the tangent to the parallel.

The Lagrange equation of the coordinate r

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{r}}-\frac{\partial T}{\partial r}=Q_r$$

takes on the form

$$m\,(\ddot{r}-r\dot{\varphi}^2-r\sin^2\varphi\dot{\psi}^2)=F_r$$

For the coordinates φ and ψ we find

$$m\,(r\ddot{\varphi}+2\dot{r}\dot{\varphi}-r\sin\varphi\cos\varphi\dot{\psi}^2)=F_\varphi$$

$$m\,(r\sin\varphi\ddot{\psi}+2\sin\varphi\dot{r}\dot{\psi}+2r\cos\varphi\dot{\varphi}\dot{\psi})=F_\psi$$

We have obtained three differential equations of the motion of a free particle in spherical coordinates.

---

* We send the tangents (to the meridian) and the parallels in the direction of appropriate increasing coordinates φ and ψ.

7. Investigating Lagrange's Equations

In order to form the Lagrange equations it is necessary first to find the expression for the kinetic energy as a function of the time t, the generalized coordinates q_i and the generalized velocities $\dot{q}_i$ ($i = 1, \ldots, n$). Let us do this in the general form:

$$T = \frac{1}{2} \sum_{\nu=1}^{N} m_\nu \dot{\boldsymbol{r}}_\nu^2 = \frac{1}{2} \sum_{\nu=1}^{N} m_\nu \left(\sum_{i=1}^{n} \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \dot{q}_i + \frac{\partial \boldsymbol{r}_\nu}{\partial t} \right)^2 =$$

$$= \frac{1}{2} \sum_{i,\, k=1}^{n} a_{ik} \dot{q}_i \dot{q}_k + \sum_{i=1}^{n} a_i \dot{q}_i + a_0 \tag{1}$$

Here, the coefficients a_{ik}, a_i, a_0 are functions of t, $q_1, \ldots, q_n$ defined by the equations

$$a_{ik} = \sum_{\nu=1}^{N} m_\nu \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \frac{\partial \boldsymbol{r}_\nu}{\partial q_k} \quad (i, k = 1, \ldots, n)^* \tag{2}$$

$$a_i = \sum_{\nu=1}^{N} m_\nu \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \frac{\partial \boldsymbol{r}_\nu}{\partial t} \quad (i = 1, \ldots, n) \tag{3}$$

$$a_0 = \frac{1}{2} \sum_{\nu=1}^{N} m_\nu \left(\frac{\partial \boldsymbol{r}_\nu}{\partial t} \right)^2 \tag{4}$$

Formula (1) shows that the kinetic energy of a holonomic system is a function (polynomial) of the second degree in the generalized velocities:

$$T = T_2 + T_1 + T_0 \tag{5}$$

where

$$T_2 = \frac{1}{2} \sum_{i,\, k=1}^{n} a_{ik} \dot{q}_i \dot{q}_k, \quad T_1 = \sum_{i=1}^{n} a_i \dot{q}_i, \quad T_0 = a_0 \tag{6}$$

As was explained in Sec. 1, in the case of a scleronomic system, the time does not explicitly enter into the relation between $\boldsymbol{r}_\nu$ and q_i, and for this reason

$$\frac{\partial \boldsymbol{r}_\nu}{\partial t} = 0 \quad (\nu = 1, \ldots, N)$$

But then, according to equations (3) and (4),

$$a_0 = 0, \; a_i = 0 \quad (i = 1, \ldots, n)$$

---

* From formulas (2) it will be seen that $a_{ik} = a_{ki}$ ($i, k = 1, \ldots, n$).

and

$$T=T_2=\frac{1}{2}\sum_{i,\,k=1}^{n} a_{ik}\dot{q}_i\dot{q}_k$$

Thus, the kinetic energy of a scleronomic system appears in the form of a homogeneous function of the second degree (quadratic form) of the generalized velocities.

It will be noted that in an arbitrary (scleronomic or rheonomic) holonomic system, the form T_2 is always nondegenerate, i.e. a determinant made up of its coefficients is different from zero:

$$\det\,(a_{ik})_{i,\,k=1}^{n}\neq 0 \tag{7}$$

Indeed, let

$$\det\,(a_{ik})_{i,\,k=1}^{n}\equiv 0$$

Then the system of homogeneous linear equations

$$\sum_{k=1}^{n} a_{ik}\lambda_k=0 \quad (i=1,\ \ldots,\ n) \tag{8}$$

has a real nonzero solution.

Multiplying the set of equations (8) termwise by λ_i, then summing with respect to i from 1 to n and utilizing formulas (2), we get

$$0=\sum_{i,\,k=1}^{n} a_{ik}\lambda_i\lambda_k=\sum_{i,\,k=1}^{n}\left(\sum_{\nu=1}^{N} m_\nu\frac{\partial \boldsymbol{r}_\nu}{\partial q_i}\cdot\frac{\partial \boldsymbol{r}_\nu}{\partial q_k}\right)\lambda_i\lambda_k=$$
$$=\sum_{\nu=1}^{N} m_\nu\left(\sum_{i=1}^{n}\lambda_i\frac{\partial \boldsymbol{r}_\nu}{\partial q_i}\right)^2$$

whence

$$\sum_{i=1}^{n}\lambda_i\frac{\partial \boldsymbol{r}_\nu}{\partial q_i}=0 \quad (\nu=1,\ \ldots,\ N) \tag{9}$$

These N vector equations may be replaced by $3N$ scalar equations:

$$\sum_{i=1}^{n}\frac{\partial x_\nu}{\partial q_i}\lambda_i=0,\quad \sum_{i=1}^{n}\frac{\partial y_\nu}{\partial q_i}\lambda_i=0,\quad \sum_{i=1}^{n}\frac{\partial z_\nu}{\partial q_i}\lambda_i=0 \quad (\nu=1,\ \ldots,\ N) \tag{9'}$$

The equations (9′) show that in the Jacobian functional matrix

$$\begin{Vmatrix} \frac{\partial x_1}{\partial q_1}, & \ldots, & \frac{\partial x_1}{\partial q_n} \\ \frac{\partial y_1}{\partial q_1}, & \ldots, & \frac{\partial y_1}{\partial q_n} \\ \frac{\partial z_1}{\partial q_1}, & \ldots, & \frac{\partial z_1}{\partial q_n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \frac{\partial x_N}{\partial q_1}, & \ldots, & \frac{\partial x_N}{\partial q_n} \\ \frac{\partial y_N}{\partial q_1}, & \ldots, & \frac{\partial y_N}{\partial q_n} \\ \frac{\partial z_N}{\partial q_1}, & \ldots, & \frac{\partial z_N}{\partial q_n} \end{Vmatrix} \tag{10}$$

the columns are linearly dependent, i.e. the rank ρ of this functional matrix is less than n. Then among the $3N$ functions $x_1, y_1, z_1, \ldots, x_N, y_N, z_N$ of the n arguments $q_1 \ldots, q_n$ (t is regarded as a parameter) there are ρ independent quantities in terms of which all the remaining Cartesian coordinates of the points of the system may be expressed. This is a contradiction, since the minimal number of independent coordinates of the system is equal to the number of degrees of freedom n, while $\rho < n$. The inequality (7) is thus established.*

The property of coefficients of quadratic form T_2 expressed by the inequality (7) is very essential and we shall repeatedly make use of it in the sequel. Note that since always $T_2 \geqslant 0$ (T_2 is the kinetic energy in the case of "frozen" constraints!), it follows from inequality (7) that the quadratic form $T_2 = \frac{1}{2}\sum\limits_{i,k=1}^{n} a_{ik}\dot{q}_i\dot{q}_k$ is positive definite, i.e. $T_2 \geqslant 0$, and $T_2 = 0$ only when all $\dot{q}_i$ ($i = 1, \ldots, n$) are equal to zero. Therefore, for the coefficients a_{ik} we have the determinantal inequalities of Sylvester:

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ \ldots, \quad \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} > 0 \tag{11}$$

Putting the expression (1) for kinetic energy into the Lagrange equations

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = Q_i \quad (i = 1, \ldots, n) \tag{12}$$

---

* The rank of the functional matrix (10) may be less than n at individual (singular) points. At these points the equality $\det (a_{ik})_{i,\,k=1}^{n} = 0$ is possible. In the sequel we shall not consider such special positions of the system.

we get

$$\sum_{k=1}^{n} a_{ik}\ddot{q}_k + (**) = Q_i(t, q_j, \dot{q}_j) \quad (i = 1, \ldots, n) \tag{13}$$

Here, the symbol (**) indicates the sum of the terms not involving second derivatives of the coordinates with respect to the time. The right-hand sides likewise do not contain second derivatives since in the general case they are functions of the quantities t, q_j, $\dot{q}_j$ ($j = 1, \ldots, n$).

Since $\det (a_{ik})_{i,k=1}^{n} \neq 0$, it follows that the equations (13) may be solved for the second derivatives and represented in the form

$$\ddot{q}_i = G_i(t, q_k, \dot{q}_k) \quad (i = 1, \ldots, n). \tag{14}$$

But then, as we know from the theory of differential equations, for certain assumptions relative to the right-hand sides G_i, which in mechanics are always assumed to be satisfied*, there is one and only one solution of the Lagrange equations for arbitrary preassigned initial q_i^0, $\dot{q}_i^0$ with $t = t_0$ ($i = 1, \ldots, n$). Thus, the motion of a holonomic system is uniquely determined by specifying the initial position (q_i^0) and the initial velocities ($\dot{q}_i^0$).

8. Theorem on Variation of Total Energy. Potential, Gyroscopic and Dissipative Forces

If the generalized forces do not depend on the generalized velocities

$$Q_i = Q_i(t, q_1, \ldots, q_n) \quad (i = 1, \ldots, n) \tag{1}$$

and there exists a function $\Pi(t, q_1, \ldots, q_n)$ such that

$$Q_i = -\frac{\partial \Pi}{\partial q_i} \quad (i = 1, \ldots, n) \tag{2}$$

then the forces Q_i are called *potential*, and the function Π is the *potential of the forces* or the *potential energy*. Equations (2), which determine the potential Π may be written as**

$$\delta A = \sum_{i=1}^{n} Q_i \delta q_i = -\delta \Pi \tag{3}$$

---

* For instance, if the functions G_i ($i = 1, \ldots, n$) have continuous first-order partial derivatives.

** The time t is first fixed when computing the virtual differential $\delta\Pi$. Therefore $\delta\Pi = \sum_{i=1}^{n} \frac{\partial \Pi}{\partial q_i} \delta q_i$.

Let us now consider the general case when in addition to the potential forces determined by the potential Π, the system is acted upon also by nonpotential forces

$$\tilde{Q}_i = \tilde{Q}_i(t, q_j, \dot{q}_j) \quad (i = 1, \ldots, n). \tag{4}$$

Then

$$Q_i = -\frac{\partial \Pi}{\partial q_i} + \tilde{Q}_i \tag{5}$$

and the Lagrange equations assume the form

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = -\frac{\partial \Pi}{\partial q_i} + \tilde{Q}_i \quad (i = 1, \ldots, n) \tag{6}$$

We now consider the total energy E, which is equal to the sum of the kinetic energy and potential energy

$$E = T + \Pi \tag{7}$$

and compute the derivative $\frac{dE}{dt}$. To do this, we first find

$$\frac{dT}{dt} = \sum_{i=1}^{n}\left(\frac{\partial T}{\partial q_i}\dot{q}_i + \frac{\partial T}{\partial \dot{q}_i}\ddot{q}_i\right) + \frac{\partial T}{\partial t} =$$

$$= \frac{d}{dt}\sum_{i=1}^{n}\frac{\partial T}{\partial \dot{q}_i}\dot{q}_i + \sum_{i=1}^{n}\left(\frac{\partial T}{\partial q_i} - \frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i}\right)\dot{q}_i + \frac{\partial T}{\partial t} \tag{8}$$

Noting that $T = T_2 + T_1 + T_0$ and utilizing the Lagrange equations (6), we obtain[*]

$$\frac{dT}{dt} = \frac{d}{dt}(2T_2 + T_1) + \frac{\partial T}{\partial t} + \sum_{i=1}^{n}\left(\frac{\partial \Pi}{\partial q_i} - \tilde{Q}_i\right)\dot{q}_i =$$

$$= 2\frac{dT}{dt} - \frac{d}{dt}(T_1 + 2T_0) + \frac{\partial T}{\partial t} + \frac{d\Pi}{dt} - \frac{\partial \Pi}{\partial t} - \sum_{i=1}^{n}\tilde{Q}_i\dot{q}_i \tag{9}$$

---

[*] The Euler formula $\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}x_i = mf$ holds for a homogeneous function $f(x_1, \ldots, x_n)$ of the mth degree. Applying this formula to the linear form T_1 and to the quadratic form T_2, we find

$$\sum_{i=1}^{n}\frac{\partial T_2}{\partial \dot{q}_i}\dot{q}_i = 2T_2, \quad \sum_{i=1}^{n}\frac{\partial T_1}{\partial \dot{q}_i}\dot{q}_i = T_1$$

The validity of these identities also follows directly from the expressions for T_2 and T_1 given on page 45.

From this, taking into account equation (7), we obtain

$$\frac{dE}{dt}=\sum_{i=1}^{n}\tilde{Q}_i\dot{q}_i+\frac{d}{dt}(T_1+2T_0)-\frac{\partial T}{\partial t}+\frac{\partial \Pi}{\partial t} \tag{10}$$

The expression on the right,

$$\sum_{i=1}^{n}\tilde{Q}_i\dot{q}_i=\frac{\sum_{i=1}^{n}\tilde{Q}_i dq_i}{dt}=\frac{\delta\tilde{A}}{dt} \tag{11}$$

where $\delta\tilde{A}$ is the elementary work of the nonpotential forces $\tilde{Q}_i$, is the *power* of the nonpotential forces $\tilde{Q}_i$ $(i=1, \ldots, n)$. The term on the right-hand side,

$$\frac{d}{dt}(T_1+2T_0)-\frac{\partial T}{\partial t} \tag{12}$$

is different from zero only for a rheonomic system (for a scleronomic system, $T_1 = T_0 = 0$, and $\frac{\partial T}{\partial t} = 0$). The latter term is nonzero only when the potential energy Π is explicitly dependent on the time.

Formula (10) determines the change in the total energy of an arbitrary holonomic system in motion. Let us consider some special cases.

(a) A scleronomic system. Then

$$\frac{dE}{dt}=\sum_{i=1}^{n}\tilde{Q}_i\dot{q}_i+\frac{\partial \Pi}{\partial t} \tag{13}$$

(b) *A* scleronomic system where the potential energy is not explicitly dependent on the time. Then

$$\frac{dE}{dt}=\sum_{i=1}^{n}\tilde{Q}_i\dot{q}_i \tag{14}$$

For such a system the derivative of the total energy with respect to the time is equal to the power of the nonpotential forces.

(c) A conservative system, i.e. (1) a scleronomic system, (2) a system where all forces are potential, and (3) where the potential energy Π is not explicitly dependent on the time. According to equation (10), for a conservative system

$$\frac{dE}{dt}=0 \tag{15}$$

i.e. for any motion of the system

$$E=\text{const}=h \tag{16}$$

The total energy of a conservative system does not change when the system is in motion.

Equation (16), which does not involve $\ddot{q}_i$ but involves the arbitrary constant h, determines the first integral of the equations of motion. Equation (16) is called the *energy integral*.

Nonpotential forces are called *gyroscopic* forces if their power is zero:

$$\sum_{i=1}^{n} \tilde{Q}_i \dot{q}_i = 0 \tag{17}$$

and *dissipative* forces if their power* is negative or zero:

$$\sum_{i=1}^{n} \tilde{Q}_i \dot{q}_i \leqslant 0 \tag{18}$$

If the potential energy is not explicitly dependent on t, then from equations (14) and (17) it follows that $\frac{dE}{dt} = 0$, and thus for a scleronomic system with gyroscopic forces we also have the energy integral

$$E = \text{const}$$

---

* In the case of a scleronomic system

$$\sum_{\nu=1}^{N} \boldsymbol{F}_\nu \, d\boldsymbol{r}_\nu = \sum_{i=1}^{n} Q_i \, dq_i$$

whence after termwise division by dt we find

$$\sum_{\nu=1}^{N} \boldsymbol{F}_\nu \boldsymbol{r}_\nu = \sum_{i=1}^{n} Q_i \dot{q}_i \tag{*}$$

For this reason equation (17) expresses the condition of gyroscopicity

$$\sum_{\nu=1}^{N} \boldsymbol{F}_\nu \boldsymbol{v}_\nu = 0$$

and equation (18) the condition of dissipativity

$$\sum_{\nu=1}^{N} \boldsymbol{F}_\nu \boldsymbol{v}_\nu \leqslant 0$$

In the case of a rheonomic system, equation (*) may not hold. Then $\delta \boldsymbol{r}_\nu = d\boldsymbol{r}_\nu - \frac{\partial \boldsymbol{r}_\nu}{\partial t} dt$ and from the equation $\sum_{\nu=1}^{N} \boldsymbol{F}_\nu \delta \boldsymbol{r}_\nu = \sum_{i=1}^{n} Q_i \delta q_i$ it follows that $\sum_{\nu=1}^{N} \boldsymbol{F}_\nu \left(\boldsymbol{v}_\nu - \frac{\partial \boldsymbol{r}_\nu}{\partial t} \right) = \sum_{i=1}^{n} Q_i \dot{q}_i$.

But if such a system is acted upon by dissipative forces, then

$$\frac{dE}{dt} \leqslant 0$$

when the system is in motion; i.e. the total energy diminishes during the motion.* In that case we call the system a *dissipative* system.

In the relations (17) and (18), the generalized forces $\tilde{Q}_i$ in the general case depend on the generalized velocities. Consider some important special cases in which this dependence is linear and homogeneous.

1°. Let

$$\tilde{Q}_i = \sum_{k=1}^{n} \gamma_{ik} \dot{q}_k \quad (i = 1, \ldots, n) \tag{19}$$

and let the matrix of the coefficients γ_{ik} be skew-symmetric**:

$$\gamma_{ik} = -\gamma_{ki} \quad (i, k = 1, \ldots, n) \tag{20}$$

Then the forces (19) are gyroscopic.

Indeed, in this case

$$\sum_{i=1}^{n} \tilde{Q}_i \dot{q}_i = \sum_{i, k=1}^{n} \gamma_{ik} \dot{q}_i \dot{q}_k = \sum_{i=1}^{n} \gamma_{ii} \dot{q}_i^2 + \sum_{i<k}^{1, \ldots, n} (\gamma_{ik} + \gamma_{ki}) \dot{q}_i \dot{q}_k = 0$$

This equation shows that the skew-symmetry of the matrix of coefficients γ_{ik} is not only a sufficient but also a necessary condition for the forces (19) applied to the scleronomic system to be gyroscopic.

Example 1. For a scleronomic system, the Coriolis forces of inertia are gyroscopic forces. Indeed, the Coriolis force of inertia applied to a particle P_ν of a system is determined from the formula

$$\boldsymbol{F}_\nu = -2m_\nu (\boldsymbol{\omega} \times \boldsymbol{v}_\nu)$$

where, m_ν is the mass of the particle P_ν, $\boldsymbol{v}_\nu$ is its velocity in the noninertial system of coordinate axes under consideration, and $\boldsymbol{\omega}$ is the angular velocity of rotation of this system relative to some inertial system of coordinates ($\nu = 1, \ldots, N$). Then

$$\sum_{\nu=1}^{N} \boldsymbol{F}_\nu \boldsymbol{v}_\nu = 0$$

Example 2. Let a rigid body with a stationary point O be acted upon by forces with principal moment $\boldsymbol{L}_0 = I(\boldsymbol{\omega}_1 \times \boldsymbol{\omega}_2)$ where I is a scalar, and let $\boldsymbol{\omega} = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_2$ be the angular velocity of the body. Then the forces applied to the body are gyroscopic since their power is equal to zero:

$$\boldsymbol{L}_0 \boldsymbol{\omega} = 0$$

---

* Energy is dissipated in the case of dissipative forces; hence the term.

** In the case of a skew-symmetric matrix $\| \gamma_{ik} \|$, always $\gamma_{ii} = 0$ $(i = 1, \ldots, n)$.

If the rigid body possesses dynamical symmetry, I is the **moment of** inertia about the axis of symmetry, ω_2 is the angular velocity of "pure rotation" directed along the axis of symmetry, and ω_1 is the angular velocity of precessional motion, then the moment $\boldsymbol{L}_0 = I\,(\omega_1 \times \omega_2)$ is called *gyroscopic*. Thus, the forces that create the gyroscopic moment are gyroscopic.

2°. Let

$$\widetilde{Q}_i = -\sum_{k=1}^{n} b_{ik}\dot{q}_k \quad (i = 1, \ldots, n) \tag{21}$$

where the matrix of coefficients b_{ik} is symmetrical,

$$b_{ik} = b_{ki} \quad (i = 1, \ldots, n) \tag{21'}$$

and let the quadratic form $\sum\limits_{i,\,k=1}^{n} b_{ik}\dot{q}_i\dot{q}_k$ be positive; i.e.

$$\sum_{i,\,k=1}^{n} b_{ik}\dot{q}_i\dot{q}_k \geqslant 0 \tag{22}$$

Then for the scleronomic system the power of the forces is equal to

$$\sum_{i=1}^{n} \widetilde{Q}_i\dot{q}_i = -\sum_{i,\,k=1}^{n} b_{ik}\dot{q}_i\dot{q}_k \leqslant 0 \tag{23}$$

and the forces $\widetilde{Q}_i$ are dissipative.

In this case the quadratic form

$$R = \frac{1}{2}\sum_{i,\,k=1}^{n} b_{ik}\dot{q}_i\dot{q}_k \tag{24}$$

is called *Rayleigh's dissipative function*. It is easy to see that the generalized forces (21) are obtained from Rayleigh's dissipative function by means of the formulas

$$\widetilde{Q}_i = -\frac{\partial R}{\partial \dot{q}_i} \quad (i = 1, \ldots, n) \tag{25}$$

If the system is scleronomic and the potential energy does not depend explicitly on the time, then by virtue of the equalities (14), (23), and (25)

$$\frac{dE}{dt} = \sum_{i=1}^{n} \widetilde{Q}_i\dot{q}_i = -2R \tag{26}$$

This formula points to the physical meaning of Rayleigh's function: the doubled Rayleigh function is equal to the rate of decrease in total energy.

If the Rayleigh function (24) is a positive definite quadratic form of the generalized velocities, then one speaks of the *total dissipation* *of energy*. In this case, we will call the system definite dissipative. According to formula (26), the total energy of such a system strictly diminishes.

By way of illustration, let us consider the forces of resistance of the medium applied to points of the system, the forces being proportional to the first power of the velocities of the points:

$$\boldsymbol{F}_\nu = -\beta \boldsymbol{v}_\nu \qquad (\nu = 1, \ldots, N) \tag{27}$$

In this case

$$\sum_{\nu=1}^{N} \boldsymbol{F}_\nu \boldsymbol{v}_\nu = -2R \tag{28}$$

where

$$R = \frac{1}{2}\beta \sum_{\nu=1}^{N} v_\nu^2 \tag{29}$$

9. Electromechanical Analogies

In this section we will show how the equations of analytical mechanics may be applied not only to mechanical systems but to electrical and electromechanical systems as well.

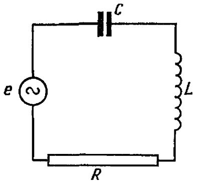

Fig. 25

Consider a circuit with inductance L, resistance R and capacitance C connected in series (Fig. 25). For these elements, the relation between the voltage u (the difference between the potentials at the ends of the element) and the current i $\left(i = \frac{dq}{dt} \text{ where } q \text{ is the charge}\right)$ will, respectively, be equal to

$$u = L\frac{di}{dt}, \quad u = Ri, \quad u = \frac{1}{C}\int i\,dt \tag{1}$$

If in addition the circuit has an external source of electromotive force $e(t)$. then we write that the electromotive force is equal to the sum of the voltages across the separate elements, and we have

$$L\frac{di}{dt}+Ri+\frac{1}{C}\int i\,dt=e(t) \tag{2}$$

or

$$L\frac{d^2q}{dt^2}+R\frac{dq}{dt}+\frac{q}{C}=e(t) \tag{3}$$

This equation is an analogue of the equation of mechanical oscillations:

$$a\frac{d^2q}{dt^2}+b\frac{dq}{dt}+cq=Q(t) \tag{4}$$

Here to the inductance L there corresponds the inertial coefficient (generalized mass) a, to the ohmic resistance R the dissipative coefficient b, to the coefficient $\frac{1}{C}$, where C is the capacitance, corresponds the reduced coefficient of elastic force c; the charge q corresponds to the generalized coordinate q, and the electromotive force $e(t)$ to the generalized force $Q(t)$.

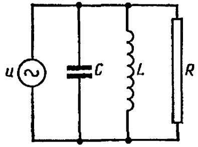

Fig. 26

On the other hand, in the circuit shown in Fig. 26 the currents passing through the inductor, the resistor and the capacitor are additive, and so

$$\frac{u}{R}+\frac{1}{L}\int u\,dt+C\frac{du}{dt}=i(t) \tag{5}$$

Differentiating termwise, we get

$$C\frac{d^2u}{dt^2}+\frac{1}{R}\frac{du}{dt}+\frac{1}{L}u=\frac{di}{dt}$$

Here we have another system of analogies in which the voltage u corresponds to the coordinate q, the mechanical coefficients a, b, c are replaced by C, $\frac{1}{R}$, $\frac{1}{L}$; and the quantity $\frac{di}{dt}$ corresponds to the generalized force $Q(t)$.

Two electric systems having the same (up to notations) equations are two different electric analogues of one and the same mechanical system.

To the kinetical and potential energies, the Rayleigh function, and the generalized force of a mechanical system with one degree of freedom

$$T=\frac{1}{2}a\dot{q}^2,\ \widetilde{R}=\frac{1}{2}b\dot{q}^2,\ \Pi=\frac{1}{2}cq^2,\ Q=Q(t)$$

in the first system of analogies there correspond the quantities

$$T=\frac{1}{2}L\dot{q}^2,\ \widetilde{R}=\frac{1}{2}R\dot{q}^2,\ \Pi=\frac{1}{2C}q^2,\ e=e(t)$$

and in the second, the quantities

$$T=\frac{1}{2}C\dot{u}^2,\ \widetilde{R}=\frac{1}{2R}\dot{u}^2,\ \Pi=\frac{1}{2L}u^2,\ \frac{di}{dt}$$

Thus, the systems of electromechanical analogies are determined by the following table:

Mechanical: q	a	b	c	Q	$T=\frac{1}{2}a\dot{q}^2$	$\widetilde{R}=\frac{1}{2}b\dot{q}^2$	$\Pi=\frac{1}{2}cq^2$
1st electrical: q	L	R	$\frac{1}{C}$	e	$\frac{1}{2}L\dot{q}^2$	$\frac{1}{2}R\dot{q}^2$	$\frac{1}{2C}q^2$
2nd electrical: u	C	$\frac{1}{R}$	$\frac{1}{L}$	$\frac{di}{dt}$	$\frac{1}{2}C\dot{u}^2$	$\frac{1}{2R}\dot{u}^2$	$\frac{1}{2L}u^2$

Let us consider the electric circuit shown in Fig. 27 as an example of a more complex illustration.

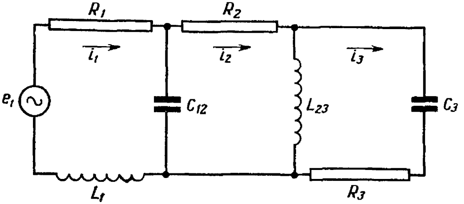

Fig. 27

We form the Lagrange equations adhering to the first system of analogies; first we compute

$$T=\frac{1}{2}L_1\dot{q}_1^2+\frac{1}{2}L_{23}(\dot{q}_2-\dot{q}_3)^2$$

$$\widetilde{R}=\frac{1}{2}R_1\dot{q}_1^2+\frac{1}{2}R_2\dot{q}_2^2+\frac{1}{2}R_3\dot{q}_3^2$$

$$\Pi=\frac{1}{2C_3}q_3^2+\frac{1}{2C_{12}}(q_1-q_2)^2$$

Also, $e_2 = e_3 = 0$. Put

$$e_1 = A \sin \Omega t$$

Now write down the Lagrange equations:

$$L_1 \ddot{q}_1 + R_1 \dot{q}_1 + \frac{1}{C_{12}} q_1 - \frac{1}{C_{12}} q_2 = A \sin \Omega t$$

$$L_{23} \ddot{q}_2 - L_{23} \ddot{q}_3 + R_2 \dot{q}_2 + \frac{1}{C_{12}} q_2 - \frac{1}{C_{12}} q_1 = 0$$

$$L_{23} \ddot{q}_3 - L_{23} \ddot{q}_2 + R_3 q_3 + \frac{1}{C_3} q_3 = 0$$

These equations will be the equations of the electric circuit depicted in Fig. 27.

10. Appell's Equations for Nonholonomic Systems. Pseudocoordinates

In this section we shall derive the Appell equations that determine the motion of a nonholonomic system. Let d finite and g differential constraints be imposed on a nonholonomic system (see Sec. 1). First utilizing only d finite constraints, we express the radii vectors of the points of the system in terms of $m = 3N - d$ independent coordinates $q_1, \ldots, q_m$ and the time t:

$$\boldsymbol{r}_\nu = \boldsymbol{r}_\nu (t, q_1, \ldots, q_m) \quad (\nu = 1, \ldots, N) \tag{1}$$

From this

$$\dot{\boldsymbol{r}}_\nu = \sum_{i=1}^{m} \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \dot{q}_i + \frac{\partial \boldsymbol{r}_\nu}{\partial t} \quad (\nu = 1, \ldots, N) \tag{2}$$

and

$$\delta \boldsymbol{r}_\nu = \sum_{i=1}^{m} \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \delta q_i \quad (\nu = 1, \ldots, N) \tag{2'}$$

However, $\boldsymbol{r}_\nu$ and $\dot{\boldsymbol{r}}_\nu$ $(\nu = 1, \ldots, N)$ also satisfy the differential constraints*

$$\sum_{\nu=1}^{N} \boldsymbol{l}_{\beta\nu} \dot{\boldsymbol{r}}_\nu + D_\beta = 0 \quad (\beta = 1, \ldots, g) \tag{3}$$

where $\boldsymbol{l}_{\beta\nu}$ and D_β are functions of t and $\boldsymbol{r}_\nu$ $(\nu = 1, \ldots, N)$.

---

* The functions (1), when substituted into the equations of finite constraints, convert the latter into identities. Therefore, when using the representation (1) it suffices to consider only the differential constraints.

Substituting expressions (1) and (2) for $\boldsymbol{r}_\nu$ and $\dot{\boldsymbol{r}}_\nu$ into the constraint equations (3) we represent these equations in the form

$$\sum_{i=1}^{m} A_{\beta i}\dot{q}_i + A_\beta = 0 \quad (\beta = 1, \ldots, g) \tag{4}$$

where the coefficients $A_{\beta i}$ of $\dot{q}_i$ and the absolute terms A_β are functions of t and $q_1, \ldots, q_m$.

Thus for a nonholonomic system the coordinates $q_1, \ldots, q_m$ can take on arbitrary values, but then the generalized velocities $\dot{q}_1, \ldots, \dot{q}_m$ can no longer be arbitrary; they are connected by the relations (4). Considering the g constraints of (4) independent, we can express g generalized velocities from equations (4), for example $\dot{q}_{n+1}, \ldots, \dot{q}_m$ in terms of the remaining $\dot{q}_1, \ldots, \dot{q}_n$ ($n = m - g = 3N - d - g$ is the number of degrees of freedom of the system; see page 15). The velocities $\dot{q}_1, \ldots, \dot{q}_n$ may be given arbitrary values and then the values of the remaining velocities will be determined.

However, we will take the more general path and for the independent quantities will take not n (n is the number of degrees of freedom) of the generalized velocities, but some n independent linear combinations of these velocities*

$$\dot{\pi}_s = \sum_{i=1}^{m} f_{si}\dot{q}_i \quad (s = 1, \ldots, n) \tag{5}$$

where f_{si} are functions of t and $q_1, \ldots, q_m$.

Only one condition need be imposed on the linear forms (5): these n linear forms together with the g linear forms

$$\sum_{i=1}^{m} A_{\beta i}\dot{q}_i \quad (\beta = 1, \ldots, g)$$

must constitute a complete system made up of $m = n + g$ linearly independent forms; i.e. the determinant composed of the coefficients of these m forms must be other than zero. Then the quantities $\dot{\pi}_s$ ($s = 1, \ldots, n$) can take on arbitrary values, since for any values of these quantities we will find the corresponding $\dot{q}_i$ ($i = 1, \ldots, m$)

---

* It will be convenient to denote the linear combinations (5) by $\dot{\pi}_s$, although the symbol π_s may be meaningless since the right-hand side of (5) may not be a total derivative.

by solving the set of linear equations (4) and (5). In this way we get

$$\dot{q}_i = \sum_{s=1}^{n} h_{is}\dot{\pi}_s + h_i \quad (i = 1, \ldots, m) \tag{6}$$

where h_{is} and h_i are functions of t and $q_1, \ldots, q_m$.

The quantities $\dot{\pi}_s$, which are linear forms of the generalized velocities, will be called *pseudovelocities*, and the symbols π_s, *pseudocoordinates* ($s = 1, \ldots, n$). In particular, $\dot{\pi}_s$ can coincide with certain generalized velocities. But in the general case of $m + n$ quantities, $\dot{\pi}_s$ and $\dot{q}_i$ are connected by the relations (5) and (6).

In order to find the restrictions imposed by differential constraints on the virtual displacements δq_i it is necessary (see Sec. 2) in equations (4) to discard the absolute terms A_β and replace the $\dot{q}_i$ by δq_i ($i = 1, \ldots, n$).
We then obtain

$$\sum_{i=1}^{m} A_{\beta i}\delta q_i = 0 \quad (\beta = 1, \ldots, g) \tag{4'}$$

In accord with equations (5) we introduce the notation*

$$\delta\pi_s = \sum_{i=1}^{m} f_{si}\delta q_i \quad (s = 1, \ldots, n) \tag{5'}$$

By assumption, forms (4′) and (5′) are linearly independent. Therefore $\delta\pi_s$ can assume arbitrary values while the corresponding δq_i will be determined from the set of equations (4′) and (5′):

$$\delta q_i = \sum_{s=1}^{n} h_{is}\delta\pi_s \quad (i = 1, \ldots, m) \tag{6'}$$

The expression for the work of elementary forces in virtual displacements may be given as

$$\delta A = \sum_{i=1}^{m} Q_i\delta q_i \tag{7}$$

where, as for a holonomic system,

$$Q_i = \sum_{\nu=1}^{N} \boldsymbol{F}_\nu \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \quad (i = 1, \ldots, m)$$

---

* In the case of a scleronomic system $\delta q_i = dq_i = \dot{q}_i\, dt$ and therefore, in accordance with formulas (5) and (5′), $\delta\pi_s = \dot{\pi}_s\, dt$.

Now, substituting into (7) the expressions (6′) in place of δq_i, we get

$$\delta A = \sum_{i=1}^{m} Q_i \sum_{s=1}^{n} h_{is}\delta\pi_s = \sum_{s=1}^{n} \Big(\sum_{i=1}^{m} h_{is}Q_i\Big)\delta\pi_s$$

i.e.

$$\delta A = \sum_{s=1}^{n} \Pi_s \delta\pi_s \tag{8}$$

where

$$\Pi_s = \sum_{i=1}^{m} h_{is}Q_i = \sum_{i=1}^{m}\sum_{\nu=1}^{N} h_{is}\frac{\partial \boldsymbol{r}_\nu}{\partial q_i}\boldsymbol{F}_\nu \quad (s = 1, \ldots, n) \tag{9}$$

We shall call the quantities Π_s *generalized forces that correspond to the pseudocoordinates* π_s $(s = 1, \ldots, n)$.

On the other hand, substituting into (2) the expressions (6) for $\dot{q}_i$ we get

$$\dot{\boldsymbol{r}}_\nu = \sum_{s=1}^{n} \boldsymbol{e}_{\nu s}\dot{\pi}_s + \boldsymbol{e}_\nu \quad (\nu = 1, \ldots, N) \tag{10}$$

where $\boldsymbol{e}_{\nu s}$ and $\boldsymbol{e}_\nu$ $(\nu = 1, \ldots, N;\ s = 1, \ldots, n)$ are certain vector functions of t and $q_1, \ldots, q_m$.

From the equations (10) we find*

$$\delta\boldsymbol{r}_\nu = \sum_{s=1}^{n} \boldsymbol{e}_{\nu s}\delta\pi_s \quad (\nu = 1, \ldots, N) \tag{11}$$

and

$$\ddot{\boldsymbol{r}}_\nu = \sum_{s=1}^{n} \boldsymbol{e}_{\nu s}\ddot{\pi}_s + \ldots \quad (\nu = 1, \ldots, N) \tag{12}$$

where on the right-hand sides of (12) we take only the terms involving *pseudoaccelerations* $\ddot{\pi}_s$ $(s = 1, \ldots, n)$.

By means of the equations (8) and (11) we write down the general equation of dynamics

$$\delta A - \sum_{\nu=1}^{N} m_\nu \ddot{\boldsymbol{r}}_\nu \delta\boldsymbol{r}_\nu = 0 \tag{13}$$

---

* The quantities $\dot{\boldsymbol{r}}_\nu$, $\dot{\pi}_s$ and $\dot{q}_i$ are connected by the relations (2), (4) and (5). Eliminating $\dot{q}_i$ from these relations we find the formulas (10). The quantities $\delta\boldsymbol{r}_\nu$, $\delta\pi_s$ and δq_i satisfy the homogeneous relations (2′), (4′) and (5′), which differ from (2), (4) and (5) solely in the absence of absolute terms. For this reason, also the formulas (11), which are the result of eliminating δq_i from (2′), (4′) and (5′), are obtained from (10) by replacing $\dot{\boldsymbol{r}}_\nu$ by $\delta\boldsymbol{r}_\nu$, $\dot{\pi}_s$ by $\delta\pi_s$ and by rejecting the absolute terms $\boldsymbol{e}_\nu$.

as follows

$$\sum_{s=1}^{n}\left(\Pi_s-\sum_{\nu=1}^{N} m_\nu \ddot{r}_\nu e_{\nu s}\right)\delta\pi_s=0 \tag{14}$$

Since $\delta\pi_s$ are quite arbitrary multipliers, it follows that

$$\sum_{\nu=1}^{N} m_\nu \ddot{r}_\nu e_{\nu s}=\Pi_s \quad (s=1, \ldots, n) \tag{15}$$

Let us now bring into the discussion the "energy of accelerations"

$$U=\frac{1}{2}\sum_{\nu=1}^{N} m_\nu \ddot{r}_\nu^2=U(t, q_i, \dot{\pi}_s, \ddot{\pi}_s) \tag{16}$$

Noting that on the basis of formulas (12)

$$e_{\nu s}=\frac{\partial \ddot{r}_\nu}{\partial \ddot{\pi}_s} \quad (\nu=1, \ldots, N;\ s=1, \ldots, n) \tag{17}$$

we can write equations (15) as follows:

$$\frac{\partial U}{\partial \ddot{\pi}_s}=\Pi_s \ (s=1, \ldots, n). \tag{18}$$

The equations (18) were first derived by Appell and are called *Appell's equations.*

These $n=3N-d-g$ differential equations, together with the g constraint equations

$$\sum_{i=1}^{m} A_{\beta i}\dot{q}_i+A_\beta=0 \ \ (\beta=1, \ldots, g) \tag{19}$$

and the n differential relations

$$\dot{\pi}_s=\sum_{i=1}^{m} f_{si}\dot{q}_i \tag{20}$$

form the system of differential equations that determine the motion of a nonholonomic system.

Let us write the Appell equation in expanded form. To do this, put expressions (12) in formula (15) instead of $\ddot{r}_\nu$. Then we get

$$\sum_{s=1}^{n} u_{\rho s}\ddot{\pi}_s+(**)=\Pi_\rho \ \ (\rho=1, \ldots, n) \tag{21}$$

where

$$\Pi_\rho = \Pi_\rho(t, q_i, \dot{\pi}_s), \quad u_{\rho s} = u_{\rho s}(t, q_i) = \sum_{\nu=1}^{N} m_\nu e_{\nu s} e_{\nu \rho}$$

$$(\rho, s = 1, \dots, n) \tag{22}$$

The symbol (* *) in equations (21) denotes the terms that do not contain pseudoaccelerations $\ddot{\pi}_s$ $(s = 1, \dots, n)$.

It may be proved that a determinant made up of the coefficients $u_{\rho s}$ is not identically equal to zero:

$$\det (u_{\rho s})_{\rho, s=1}^{n} \neq 0^{*} \tag{23}$$

Then equations (21) may be solved for the pseudoaccelerations

$$\ddot{\pi}_s = H_s(t, q_i, \dot{\pi}_\rho) \quad (s = 1, \dots, n) \tag{24}$$

On the other hand, the relations (19) and (20) can also be represented in a form solved for $\dot{q}_i$ $(i = 1, \dots, m)$ [see formulas (6)].

Thus, the motion of a nonholonomic system is determined by a system of $n + m$ differential equations of the first order in the unknown functions $q_1, \dots, q_m, \dot{\pi}_1, \dots, \dot{\pi}_n$, and these equations are solved for the derivatives. Then specification of the initial data, $q_1^0, \dots, q_m^0, \dot{\pi}_1^0, \dots, \dot{\pi}_n^0$ uniquely determines the motion of the system. But with the aid of these initial data, formulas (1) and (6) are used to specify the arbitrary initial velocities and arbitrary initial position that are compatible with constraints. For this reason, specification of the initial position of the system and the initial velocities that do not contradict the finite and differential constraints uniquely determines the motion of a nonholonomic system.

Note 1. If in a special case, n independent generalized velocities, for example, $\dot{q}_1, \dots, \dot{q}_n$, are taken for the pseudovelocities, then in order to determine the appropriate generalized forces $Q_1^*, \dots, Q_n^*$ it is necessary in (7) to express $\delta q_{n+1}, \dots, \delta q_m$ in terms of $\delta q_1, \dots, \delta q_n$:

$$\delta A = \sum_{i=1}^{m} Q_i \delta q_i = \sum_{s=1}^{n} Q_s^* \delta q_s \tag{25}$$

In this case the energy of accelerations, U, may be given in the form of the function $B(t, q_1, \dots, q_m, \dot{q}_1, \dots, \dot{q}_n, \ddot{q}_1, \dots, \ddot{q}_n)$ and

---

[*] This determinant may be zero at certain points. These singular points are not considered here. Inequality (23) is substantiated in a similar manner to the inequality $\det (a_{ik})_{i, k=1}^{n} \neq 0$ on pp. 46-47.

the Appell equations take the form

$$\frac{\partial U}{\partial \ddot{q}_s} = Q_s^* \quad (s = 1, \ldots, n) \tag{26}$$

Note 2. In particular the Appell equations may also be applied to a holonomic system. In this case, all velocities $\dot{q}_i$ are independent, $Q_i = Q_i^*$ $(i = 1, \ldots, n)$, and, equations (26) are simply another notation for the Lagrange equations of the second kind.*

Example 1. By means of the Appell equations we determine the motion of the system described in the example in Sec. 3 (see page 23). This will enable the reader to compare two methods of finding the motion of a nonholonomic system—by means of Lagrange multipliers and with the aid of Appell's equations—and to be convinced of the advantages of the latter. For the independent coordinates, we introduce the coordinates x, y of the centre of the rod and the angle φ formed by the segment M_1M_2 with the horizontal axis x (Fig. 28). Then

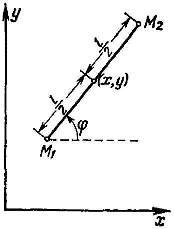

Fig. 28

$$x_1 = x - \frac{l}{2}\cos\varphi, \quad x_2 = x + \frac{l}{2}\cos\varphi$$

$$y_1 = y - \frac{l}{2}\sin\varphi, \quad y_2 = y + \frac{l}{2}\sin\varphi$$

The equation of differential constraint in the new coordinates becomes

$$\frac{\dot{x}}{\cos\varphi} = \frac{\dot{y}}{\sin\varphi}$$

It may readily be verified that the energy of accelerations, U, can be expressed in the following manner:

$$U = \frac{1}{2}(\ddot{x}_1^2 + \ddot{y}_1^2 + \ddot{x}_2^2 + \ddot{y}_2^2) = \ddot{x}^2 + \ddot{y}^2 + \frac{1}{4}l^2(\ddot{\varphi}^2 + \dot{\varphi}^4)$$

---

* However, as applied to a holonomic system, Appell's equations in pseudocoordinates yield quite different forms of the equations of motion.

We introduce the pseudovelocity $\dot{\pi}$, assuming

$$\dot{x}=\dot{\pi}\cos\varphi, \quad \dot{y}=\dot{\pi}\sin\varphi$$

then

$$U=\ddot{\pi}^2+\frac{1}{}\, l^2\ddot{\varphi}^2+\ldots$$

where the nonwritten terms do not involve accelerations. We now determine the generalized forces. To do this, write

$$\delta A=\Pi\delta\pi+\Phi\delta\varphi=-2g\delta y=-2g\sin\varphi\delta\pi$$

From this

$$\Pi=-2g\sin\varphi, \quad \Phi=0$$

Now form the Appell equations

$$\frac{\partial U}{\partial\ddot{\pi}}=\Pi, \quad \frac{\partial U}{\partial\ddot{\varphi}}=\Phi$$

In this case, these equations do not contain the coordinates x, y and have the form

$$\ddot{\pi}=-g\sin\varphi, \quad \ddot{\varphi}=0$$

Integrating, we obtain

$$\varphi=\alpha t+\beta$$

$$\frac{d\dot{\pi}}{d\varphi}=\frac{1}{\alpha}\ddot{\pi}=-\frac{g}{\alpha}\sin\varphi, \quad \dot{\pi}=\frac{g}{\alpha}\cos\varphi+\gamma$$

We find x and y:

$$\frac{dx}{d\varphi}=\frac{1}{\alpha}\dot{x}=\frac{1}{\alpha}\dot{\pi}\cos\varphi=\frac{g}{\alpha^2}\cos^2\varphi+\frac{\gamma}{\alpha}\cos\varphi$$

$$\frac{dy}{d\varphi}=\frac{1}{\alpha}\dot{y}=\frac{1}{\alpha}\dot{\pi}\sin\varphi=\frac{g}{\alpha^2}\cos\varphi\sin\varphi+\frac{\gamma}{\alpha}\sin\varphi$$

From this we have

$$x=\frac{g}{2\alpha^2}\varphi+\left(\frac{\gamma}{\alpha}+\frac{g}{2\alpha^2}\cos\varphi\right)\sin\varphi+\delta$$

$$y=-\left(\frac{\gamma}{\alpha}+\frac{g}{2\alpha^2}\cos\varphi\right)\cos\varphi+\varepsilon$$

Substituting $\alpha t+\beta$ for φ, we get the final equations of motion that contain five arbitrary constants, α, β, γ, δ and ε:

$$x=\frac{g}{2\alpha^2}(\alpha t+\beta)+\left[\frac{\gamma}{\alpha}+\frac{g}{2\alpha^2}\cos(\alpha t+\beta)\right]\sin(\alpha t+\beta)+\delta$$

$$y=-\left[\frac{\gamma}{\alpha}+\frac{g}{2\alpha^2}\cos(\alpha t+\beta)\right]\cos(\alpha t+\beta)+\varepsilon, \quad \varphi=\alpha t+\beta$$

Example 2. We will show how it is possible, from the Appell equations, to obtain the dynamical equations of Euler for a rigid body with fixed point O.

Let p, q, r be the projections of the angular velocity ω on the principal axes of inertia $O\xi$, $O\eta$, $O\zeta$. They are known to be linear combinations of the generalized velocities $\dot{\psi}$, $\dot{\theta}$, $\dot{\varphi}$, where ψ, θ, φ are the Euler angles (see pp. 35-36)*. We can therefore take $\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{r}$ for the three pseudovelocities. Let us calculate the energy of accelerations:**

$$\begin{aligned} 2U &= \int \boldsymbol{w}^2\, dm = \int (\boldsymbol{\varepsilon}\times\boldsymbol{r}+\boldsymbol{\omega}\times\boldsymbol{v})^2\, dm = \\ &= \int (\boldsymbol{\varepsilon}\times\boldsymbol{r})^2\, dm + 2\int (\boldsymbol{\varepsilon}\times\boldsymbol{r})(\boldsymbol{\omega}\times\boldsymbol{v})\, dm + \ldots = \\ &= \int (\boldsymbol{\varepsilon}\times\boldsymbol{r})^2\, dm + 2\boldsymbol{\varepsilon}\int [\boldsymbol{r}\times(\boldsymbol{\omega}\times\boldsymbol{v})]\, dm + \ldots = \\ &= \int (\boldsymbol{\varepsilon}\times\boldsymbol{r})^2\, dm + 2\boldsymbol{\varepsilon}\left[\boldsymbol{\omega}\times\int \boldsymbol{r}\times\boldsymbol{v}\, dm\right] + \ldots \end{aligned} \tag{27}$$

It will be noted that $\boldsymbol{\varepsilon}=\frac{d\boldsymbol{\omega}}{dt}=\frac{\delta\boldsymbol{\omega}}{dt}+\boldsymbol{\omega}\times\boldsymbol{\omega}=\frac{\delta\boldsymbol{\omega}}{dt}$. Here, $\frac{d}{dt}$ and $\frac{\delta}{dt}$ denote, respectively, the differentiation in a stationary system of axes and in a system of axes invariably fixed in a body. Therefore $\dot{p}$, $\dot{q}$ and $\dot{r}$ are the projections of the angular acceleration $\boldsymbol{\varepsilon}$ on the axes $O\xi$, $O\eta$, $O\zeta$.

Then, by analogy with the expression for kinetic energy,

$$2T=\int (\boldsymbol{\omega}\times\boldsymbol{r})^2\, dm = Ap^2+Bq^2+Cr^2$$

(A, B and C are the moments of inertia about the principal axes of inertia $O\xi$, $O\eta$, $O\zeta$) we can write

$$\int (\boldsymbol{\varepsilon}\times\boldsymbol{r})^2\, dm = A\dot{p}^2+B\dot{q}^2+C\dot{r}^2$$

On the other hand, the kinetic moment $\boldsymbol{G}=\int \boldsymbol{r}\times\boldsymbol{v}\, dm$ has the components Ap, Bq, Cr. For this reason, we finally obtain the following expression for $2U$:

$$2U=A\dot{p}^2+B\dot{q}^2+C\dot{r}^2+2\left[(C-B)\, qr\dot{p}+(A-C)\, rp\dot{q}+(B-A)\, pq\dot{r}\right]+\ldots$$

On the other hand, for the elementary work of external forces we have

$$\delta A = L_0\boldsymbol{\omega}\, dt = L_\xi p\, dt + L_\eta q\, dt + L_\zeta r\, dt$$

Therefore, Appell's equations yield Euler's equations directly:

$$\begin{aligned} A\frac{dp}{dt}+(C-B)\, qr &= L_\xi \\ B\frac{dq}{dt}+(A-C)\, rp &= L_\eta \\ C\frac{dr}{dt}+(B-A)\, pq &= L_\zeta. \end{aligned}$$

---

* We get expressions for p, q, r by projecting termwise on the coordinate axes the vector equation $\boldsymbol{\omega}=\boldsymbol{\omega}_\psi+\boldsymbol{\omega}_\theta+\boldsymbol{\omega}_\varphi$ where $\omega_\psi=\dot{\psi}$, $\omega_\theta=\dot{\theta}$, $\omega_\varphi=\dot{\varphi}$.

** Here we take advantage of the familiar identity

$$\boldsymbol{r}\times(\boldsymbol{\omega}\times\boldsymbol{v})+\boldsymbol{\omega}\times(\boldsymbol{v}\times\boldsymbol{r})+\boldsymbol{v}\times(\boldsymbol{r}\times\boldsymbol{\omega})=0$$

the last term on the left is zero since $\boldsymbol{v}=\boldsymbol{\omega}\times\boldsymbol{r}$. The nonwritten terms in formula (27) do not contain the angular acceleration $\boldsymbol{\varepsilon}$.

CHAPTER 2

The Equations of Motion in a Potential Field

11. Lagrange's Equations for Potential Forces. The Generalized Potential. Nonnatural Systems

Let the generalized forces Q_i be potential, i.e. let there exist a force potential (potential energy) $\Pi = \Pi(t, q_i)$ (see Sec. 8) and

$$Q_i = -\frac{\partial \Pi}{\partial q_i} \quad (i = 1, \ldots, n) \tag{1}$$

Then the Lagrange equations

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = -\frac{\partial \Pi}{\partial q_i} \quad (i = 1, \ldots, n)$$

are written in the form*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad (i = 1, \ldots, n) \tag{2}$$

where

$$L = T - \Pi \tag{3}$$

The function L is called the *Lagrangian function* or the *kinetic potential.*

The kinetic potential L, like the kinetic energy T, is a quadratic function of the generalized velocities:

$$L = L_2 + L_1 + L_0 \tag{4}$$

where

$$L_2 = \frac{1}{2}\sum_{i,\,k=1}^{n} c_{ik}\dot{q}_i\dot{q}_k, \quad L_1 = \sum_{i=1}^{n} c_i\dot{q}_i, \quad L_0 = c_0 \tag{4'}$$

Here, the coefficients c_{ik}, c_i, c_0 are functions of the coordinates $q_1, \ldots, q_n$ and the time t $(i, k = 1, \ldots, n)$. A comparison of the formula (3) and formula (5) on page 45 yields

$$L_2 = T_2, \quad L_1 = T_1, \quad L_0 = T_0 - \Pi \tag{5}$$

---

* Since the potential energy Π does not depend on the generalized velocities, and $L = T - \Pi$, it follows that $\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i}$ and $\frac{\partial L}{\partial q_i} = \frac{\partial T}{\partial q_i} - \frac{\partial \Pi}{\partial q_i}$ $(i = 1, \ldots, n)$.

Note that when the effective forces $\boldsymbol{F}_\nu = X_\nu \boldsymbol{i} + Y_\nu \boldsymbol{j} + Z_\nu \boldsymbol{k}$ $(\nu = 1, \ldots, N)$ acting on the particles have a potential $\Pi(t, x_\nu, y_\nu, z_\nu)$ in Cartesian coordinates x_ν, y_ν, z_ν $(\nu = 1, \ldots, N)$, i.e.

$$X_\nu = -\frac{\partial \Pi}{\partial x_\nu}, \quad Y_\nu = -\frac{\partial \Pi}{\partial y_\nu}, \quad Z_\nu = -\frac{\partial \Pi}{\partial z_\nu} \quad (\nu = 1, \ldots, N)$$

these forces have a potential even in the independent coordinates q_i $(i = 1, \ldots, n)$ (the converse assertion in the general case is untrue!), and this potential is the same potential Π only expressed in terms of the coordinates $q_1, \ldots, q_n$ and the time t. Indeed,

$$\sum_{i=1}^{n} Q_i \delta q_i = \sum_{\nu=1}^{N} (X_\nu \delta x_\nu + Y_\nu \delta y_\nu + Z_\nu \delta z_\nu) = -\delta \Pi = -\sum_{i=1}^{n} \frac{\partial \Pi}{\partial q_i} \delta q_i$$

whence equations (1) follow.

Let us now consider the case when in place of the ordinary potential $\Pi(t, q_k)$ there exists a *generalized potential* $V(t, q_k, \dot{q}_k)$ in terms of which the generalized forces Q_i are expressed by means of the formulas

$$Q_i = \frac{d}{dt} \frac{\partial V}{\partial \dot{q}_i} - \frac{\partial V}{\partial q_i} \quad (i = 1, \ldots, n) \tag{6}$$

Then the Lagrange equations

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = \frac{d}{dt} \frac{\partial V}{\partial \dot{q}_i} - \frac{\partial V}{\partial q_i} \quad (i = 1, \ldots, n)$$

will again be written in the form (2) where we now have

$$L = T - V \tag{7}$$

From formulas (6) it follows that

$$Q_i = \sum_{k=1}^{n} \frac{\partial^2 V}{\partial \dot{q}_i \, \partial \dot{q}_k} \ddot{q}_k + (**) \quad (i = 1, \ldots, n) \tag{8}$$

where $(**)$ denotes the sum of the terms that do not contain generalized accelerations $\ddot{q}_k$ $(k = 1, \ldots, n)$.

Inasmuch as in mechanics we consider only the case when the generalized forces Q_i are not explicitly dependent on the generalized accelerations but depend solely on the time, on the coordinates and on the generalized velocities

$$Q_i = Q_i(t, q_k, \dot{q}_k) \quad (i = 1, \ldots, n) \tag{9}$$

it then follows according to formulas (8) that all partial second-order derivatives of V with respect to the generalized velocities must be identically equal to zero, i.e. the *generalized potential V depends linearly on the generalized velocities*:

$$V = \sum_{i=1}^{n} \Pi_i \dot{q}_i + \Pi = V_1 + \Pi \tag{10}$$

where Π_i $(i = 1, \ldots, n)$ and Π are functions of the coordinates $q_1, \ldots, q_n$ and of the time t. But then, according to (7), L will again be a quadratic function in the velocities $\dot{q}_i$ and in place of equations (5) we will have*

$$L_2 = T_2, \quad L_1 = T_1 - V_1, \quad L_0 = T_0 - \Pi \tag{11}$$

Substituting expression (10) for V into formula (6), we get

$$Q_i = \frac{d\Pi_i}{dt} - \frac{\partial}{\partial q_i}\left[\sum_{k=1}^{n} \Pi_k \dot{q}_k + \Pi\right] =$$

$$= -\frac{\partial \Pi}{\partial q_i} + \sum_{k=1}^{n} \left(\frac{\partial \Pi_i}{\partial q_k} - \frac{\partial \Pi_k}{\partial q_i}\right) \dot{q}_k + \frac{\partial \Pi_i}{\partial t} \tag{12}$$

The formulas (12) show that when the linear part V_1 of the generalized potential does not depend explicitly on the time t $\left[\frac{\partial \Pi_i}{\partial t} = 0 \ (i = 1, \ldots, n)\right]$, the generalized forces Q_i are made up out of potential forces $-\frac{\partial \Pi}{\partial q_i}$ $(i = 1, \ldots, n)$ and gyroscopic forces

$$\tilde{Q}_i = \sum_{k=1}^{n} \gamma_{ik} \dot{q}_k \quad (i = 1, \ldots, n) \tag{13}$$

where

$$\gamma_{ik} = -\gamma_{ki} = \frac{\partial \Pi_i}{\partial q_k} - \frac{\partial \Pi_k}{\partial q_i} \quad (i, k = 1, \ldots, n) \tag{13'}$$

The importance of considering the generalized potential is evident from the following example.

Example. A point electric charge in an electromagnetic field is acted upon by a *Lorentz force*

$$\boldsymbol{F} = e\left[\boldsymbol{E} + \frac{\boldsymbol{v}}{c} \times \boldsymbol{H}\right] \tag{14}$$

where $\boldsymbol{v}$ is the velocity of the point, e is the charge, c is the speed of light, and $\boldsymbol{E}$ and $\boldsymbol{H}$ are the intensities of the electric and magnetic fields. The vectors $\boldsymbol{E}$ and $\boldsymbol{H}$ are expressed in terms of the scalar potential φ and the vector potential A by means of the formulas*

---

* The coefficients in the expressions for L and T are interrelated. Indeed, for the ordinary potential, $c_i = a_i$, while for the generalized potential $c_i = a_i - \Pi_i$ $(i = 1, \ldots, n)$. In both cases, $c_{ik} = a_{ik}$ $(i, k = 1, \ldots, n)$, $c_0 = a_0 - \Pi$, and $L_2 = T_2$ is a positive definite quadratic form.

$$\boldsymbol{E} = -\operatorname{grad}\varphi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}, \quad \boldsymbol{H} = \operatorname{rot} \boldsymbol{A} \tag{15}$$

We shall find the generalized potential V for the Lorentz force $\boldsymbol{F}$.

From the formulas (14) and (15) we have

$$\begin{aligned}\boldsymbol{F} &= -e\operatorname{grad}\varphi - \frac{e}{c}\frac{\partial \boldsymbol{A}}{\partial t} + \frac{e}{c}(\boldsymbol{v}\times\operatorname{rot}\boldsymbol{A}) = \\ &= -e\operatorname{grad}\varphi - \frac{e}{c}\frac{\partial \boldsymbol{A}}{\partial t} + \frac{e}{c}[(\boldsymbol{v}\nabla)\boldsymbol{A} + \boldsymbol{v}\times\operatorname{rot}\boldsymbol{A}] = \\ &= -e\operatorname{grad}\varphi - \frac{e}{c}\frac{\partial \boldsymbol{A}}{\partial t} + \frac{e}{c}\operatorname{grad}(\boldsymbol{v}\boldsymbol{A})\end{aligned} \tag{16}$$

where the velocity $\boldsymbol{v}$ in the expression grad $(\boldsymbol{v}\boldsymbol{A})$ is considered a vector that is independent of a field point**.

Hence, choosing for the independent coordinates Cartesian coordinates of the point x, y, z and setting

$$V = e\varphi - \frac{e}{c}(\boldsymbol{v}\boldsymbol{A}) \tag{17}$$

i.e.

$$V = e\varphi - \frac{e}{c}(\dot{x}A_x + \dot{y}A_y + \dot{z}A_z)$$

we have

$$F_x = -\frac{e}{c}\frac{dA_x}{dt} - \frac{\partial V}{\partial x} = \frac{d}{dt}\frac{\partial V}{\partial \dot{x}} - \frac{\partial V}{\partial x}$$

---

* See, for example, L. D. Landau and E. M. Lifshits, *Field Theory*, Moscow-Leningrad, 1948, p. 55 (In Russian).

** Here, for the expression $(v\nabla)\boldsymbol{A} = v_x\frac{\partial A}{\partial x} + v_y\frac{\partial A}{\partial y} + v_z\frac{\partial A}{\partial z}$ we take advantage of the familiar formula of vector analysis $(v\nabla)\boldsymbol{A} + \boldsymbol{v}\times\operatorname{rot}\boldsymbol{A} = \operatorname{grad}(\boldsymbol{v}\boldsymbol{A})$, in which $\boldsymbol{v}$ is regarded as a constant vector. We are readily convinced of the validity of this formula by comparing the projections on the x-, y-, and z-axis of the left and right sides of the equation. Indeed, for the x-axis,

$$\begin{aligned}&v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_x}{\partial y} + v_z\frac{\partial A_x}{\partial z} + \\ &+ v_y\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) - v_z\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) = \\ &= v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_y}{\partial x} + v_z\frac{\partial A_z}{\partial x} = \frac{\partial}{\partial x}(\boldsymbol{v}\boldsymbol{A})\end{aligned}$$

Analogous formulas hold for the projections on the y-axis and the z-axis.

Similar formulas are found for F_y and F_z. Thus, the generalized potential of the Lorentz force (14) is determined from the formula (17). For the Lagrangian function L we have the expression

$$L=T-V=\frac{1}{2}mv^2-e\varphi+\frac{e}{c}(vA) \tag{18}$$

Classical systems in which the forces have the ordinary potential $\Pi(t, q_i)$ or the generalized potential $V(t, q_i, \dot{q}_i)$ will be called *natural*. For such systems the Lagrangian function L is a quadratic function of the generalized velocities, i.e. it is given by the expression (4), where L_2 is a positive definite quadratic form in the generalized velocities.

As an example of a nonnatural system we can examine, in relativistic theory, the motion of a particle in the absence of a force field. The motion of such a particle is determined by the Lagrange equations in which

$$L=-mc^2\left(1-\frac{v^2}{c^2}\right)^{1/2}$$

where $v^2=\dot{x}^2+\dot{y}^2+\dot{z}^2$, and c is the speed of light. Here, L is no longer a quadratic function of the velocities $\dot{x}$, $\dot{y}$, $\dot{z}$.

If in the expression for the function L we expand $\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}$ in a power series of $\frac{v}{c}$ and reject all terms of second and higher orders in $\frac{v}{c}$, i.e. set $\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}\approx 1-\frac{1}{2}\frac{v^2}{c^2}$, then we have the "classical" expression for the Lagrangian function for an isolated particle, namely:

$$L=\frac{1}{2}mv^2+\text{const}$$

In this and the following chapters we will confine our discussion to systems of the general type,* whose motion is determined by Lagrange's equations (2) with an arbitrary function $L=L(t, q_i, \dot{q}_i)$. We shall only assume that the Hessian of the function L with respect to the generalized velocities is not identically equal to zero;**

$$\det\left(\frac{\partial^2 L}{\partial\dot{q}_i\,\partial\dot{q}_k}\right)_{i,\,k=1}^{n}\neq 0 \tag{19}$$

---

* Such propositions as hold only for natural systems will be specially stipulated.

** For natural systems, $\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_k}=\frac{\partial^2 T}{\partial\dot{q}_i\partial\dot{q}_k}=a_{ik}\ (i,\ k=1,\ \ldots,\ n)$ and hence, by what was proved in Sec. 7 (pages 46-47), inequality (19) holds true.

In expanded form, the equations (2) may be written as

$$\sum_{k=1}^{n} \frac{\partial^2 L}{\partial \dot{q}_i \, \partial \dot{q}_k} \ddot{q}_k + (**) = 0 \tag{20}$$

where (**) denotes the sum of the terms that do not contain the generalized accelerations $\ddot{q}_i$ $(i = 1, \ldots, n)$. Insofar as the determinant of the system of linear (in $\ddot{q}_k$) equations (20) is different from zero [see inequality (19)], it follows that the system (20) may be solved for the generalized accelerations and written in the form

$$\ddot{q}_i = G_i (t, q_k, \dot{q}_k) \quad (i = 1, \ldots, n)$$

Therefore the conclusion drawn in Sec. 7 as to the unique determination of the motion of a system by specification of the initial data $q_i^0, \dot{q}_i^0$ $(i = 1, \ldots, n)$ holds not only for natural systems but also for systems of a more general type that are here under consideration.

12. Canonical Equations of Hamilton

Lagrange showed how the differential equations of motion of a system are written down if the kinetical potential (the Lagrangian function) $L = L(t, q_i, \dot{q}_i)$ is known.

We will call the variables $t, q_i, \dot{q}_i$ $(i = 1, \ldots, n)$, in terms of which the Lagrangian function is expressed, *Lagrangian variables*. The system of these variables characterizes the instant of time and the corresponding state of the system, i.e. the position of the system and the velocities of its particles. As has already been noted at the end of the preceding section, specification of the Lagrangian function and the initial state uniquely determines the motion of the system.

For the basic variables that characterize the state of a system, Hamilton proposed the quantities t, q_i, p_i $(i = 1, \ldots, n)$, where p_i $(i = 1, \ldots, n)$ are the *generalized momenta* defined by the equalities

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad (i = 1, \ldots, n) \tag{1}$$

We shall call the variables t, q_i, p_i $(i = 1, \ldots, n)$ *Hamiltonian variables*.

Since the Jacobian of the right sides of equations (1) with respect to the variables $\dot{q}_i$ is the Hessian (different from zero) of the function L [see the condition (19) on page 70], it follows that equations (1) may be solved for $\dot{q}_i$ $(i = 1, \ldots, n)$:

$$\dot{q}_i = \Phi_i(t, q_k, p_k) \quad (i = 1, \ldots, n) \tag{2}$$

Thus, Hamilton's variables may be expressed in terms of the Lagrange variables and vice versa, and the state of the system may be described both as a system of values of the Lagrange variables and as a system of values of Hamiltonian variables.

In the case of a natural system, L is a quadratic function (see pages 66-68) of the generalized velocities and, according to equations (1), the generalized momenta are linearly expressed in terms of the generalized velocities:

$$p_i = \sum_{k=1}^{n} a_{ik}\dot{q}_k + c_i \quad (i = 1, \ldots, n) \tag{3}$$

Solving this set of linear equations for $\dot{q}_i$*, we again get linear expressions for $\dot{q}_i$:

$$\dot{q}_i = \sum_{k=1}^{n} b_{ik}p_k + b_i \quad (i = 1, \ldots, n), \tag{4}$$

where b_{ik} and b_i are functions of $t, q_1, \ldots, q_n$.

If in a natural system the forces Q_i $(i = 1, \ldots, n)$ have an ordinary potential $\Pi(t, q_i)$, it follows from the equation $L = T - \Pi$ that**

$$p_i = \frac{\partial T}{\partial \dot{q}_i} \tag{5}$$

It will be noted that any function of Lagrangian variables

$$F = F(t, q_i, \dot{q}_i)$$

after substitution of the expressions (2) or (4) into it in place of the generalized velocities $\dot{q}_i$, is converted into a certain function $\widehat{F}(t, q_i, p_i)$ of the Hamiltonian variables. We shall call the function $\widehat{F}(t, q_i, p_i)$ the associated expression of the function $F(t, q_i, \dot{q}_i)$.

---

* As was established in Sec. 7, $\det (a_{ik})_{i,k=1}^{n} \neq 0$.

** Formulas (5) do not hold for forces that have a generalized potential. In this case [see equations (7) and (10) on pp. 67-68]

$$p_i = \frac{\partial T}{\partial \dot{q}_i} - \Pi_i \quad (i = 1, \ldots, n).$$

Example. For a free particle the Cartesian coordinates x, y, z are independent, and in the potential field $\Pi=\Pi(t, x, y, z)$ the Lagrangian function has the form

$$L=\frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2)-\Pi(t, x, y, z)$$

To the Cartesian coordinates there correspond the momenta

$$p_x=\frac{\partial L}{\partial \dot{x}}=m\dot{x}, \quad p_y=m\dot{y}, \quad p_z=m\dot{z} \tag{6}$$

If from this we determine $\dot{x}$, $\dot{y}$, $\dot{z}$ and put the expressions obtained into L, we will have the associated expression for L:

$$\widehat{L}=\frac{1}{2m}(p_x^2+p_y^2+p_z^2)-\Pi(t, x, y, z) \tag{7}$$

If in place of $\Pi(t, x, y, z)$ we have the generalized potential

$$V=\Pi_1\dot{x}+\Pi_2\dot{y}+\Pi_3\dot{z}+\Pi,$$

where Π_1, Π_2, Π_3, and Π are functions of t, x, y, and z, then from the equation

$$L=\frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2)-\Pi_1\dot{x}-\Pi_2\dot{y}-\Pi_3\dot{z}-\Pi$$

we find

$$p_x=\frac{\partial L}{\partial \dot{x}}=m\dot{x}-\Pi_1, \quad p_y=m\dot{y}-\Pi_2, \quad p_z=m\dot{z}-\Pi_3 \tag{6'}$$

and the associated expression $\widehat{L}$ has the form

$$\widehat{L}=\frac{1}{2m}(p_x^2+p_y^2+p_z^2)-\frac{1}{2m}(\Pi_1^2+\Pi_2^2+\Pi_3^2)-\Pi \tag{7'}$$

Hamilton introduced the function $H(t, q_i, p_i)$ defined by the equation

$$H=\sum_{i=1}^{n} p_i\hat{\dot{q}}_i-\widehat{L}, \tag{8}$$

and demonstrated that with the aid of this function the equations of motion may be written in the form of the following system of $2n$ ordinary differential equations of the first order:

$$\frac{dq_i}{dt}=\frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q_i} \quad (i=1, \ldots, n) \tag{9}$$

These equations are called *canonical equations* or *Hamilton's equations**. The function $H(t, q_i, p_i)$ defined by equation (8) is called the *Hamiltonian function.*

---

* These equations were first obtained in general form by W. R. Hamilton in 1834.

The derivation of the canonical equations of Hamilton will rest on the following mathematical theorem.

Donkin's theorem.* Given a certain function $X(x_1, \ldots, x_n)$ the Hessian of which is different from zero:

$$\det\left(\frac{\partial^2 X}{\partial x_i\,\partial x_k}\right)_{i,\,k=1}^{n} \neq 0 \tag{10}$$

Let there also be a transformation of the variables "generated" by the function $X(x_1, \ldots, x_n)$:

$$y_i = \frac{\partial X}{\partial x_i} \quad (i = 1, \ldots, n) \tag{11}$$

Then there exists a transformation, the inverse of the transformation (11), which likewise generates some function $Y(y_1, \ldots, y_n)$:

$$x_i = \frac{\partial Y}{\partial y_i} \quad (i = 1, \ldots, n) \tag{12}$$

Here, the generating function Y of the inverse transformation is connected with the generating function X of the direct transformation by the formula**

$$Y = \sum_{i=1}^{n} x_i y_i - X \tag{13}$$

If the function X contains the parameters $\alpha_1, \ldots, \alpha_m$, i.e. $X = X(x_1, \ldots, x_n;\ \alpha_1, \ldots, \alpha_m)$, then Y also contains these parameters, i.e. $Y = Y(y_1, \ldots, y_n;\ \alpha_1, \ldots, \alpha_m)$, and

$$\frac{\partial Y}{\partial \alpha_j} = -\frac{\partial X}{\partial \alpha_j} \quad (j = 1, \ldots, m) \tag{14}$$

Proof. The Hessian of the function X coincides with the Jacobian of the right-hand sides of equations (11). For this reason, the condition (10) shows that using equations (11) it is possible to express the variables $x_1, \ldots, x_n$ in terms of $y_1, \ldots, y_n$:

$$x_i = f_i(y_1, \ldots, y_n) \quad (i = 1, \ldots, n) \tag{15}$$

Let the function $Y(y_1, \ldots, y_n)$ be defined by the formula (13), in which the variables x_i are replaced by the expressions (15). Then

$$\frac{\partial Y}{\partial y_i} = \frac{\partial}{\partial y_i}\left(\sum_{k=1}^{n} x_k y_k - X\right) = \sum_{k=1}^{n} \frac{\partial x_k}{\partial y_i} y_k + x_i - \sum_{k=1}^{n} \frac{\partial X}{\partial x_k}\frac{\partial x_k}{\partial y_i}$$

---

[*] Philosoph. Trans., 1854. The transition from the variables x_i to the variables y_i $(i = 1, \ldots, n)$ which is dealt with in Donkin's theorem, is frequently called a Legendre transformation.

[**] It is assumed that on the left-hand side of (13) all the x_i are expressed in terms of y_i, that is that $Y = Y(y_1, \ldots, y_n)$.

But, by equations (11), the two sums on the right-hand side of this equation cancel and hence, the formulas (12) hold true.

Now let X contain parameters $\alpha_1, \ldots, \alpha_m$ in addition to the variables $x_1, \ldots, x_n$. Then these parameters occur in the direct transformation (11) and, consequently, in the reverse one as well:

$$x_i = f_i(y_1, \ldots, y_n;\ \alpha_1, \ldots, \alpha_m) \quad (i = 1, \ldots, n)$$

The function Y is determined by equation (13) in which the x_i are replaced by $f_i(y_1, \ldots, y_n;\ \alpha_1, \ldots, \alpha_n)$ and so*

$$\frac{\partial Y}{\partial \alpha_j} = \frac{\partial}{\partial \alpha_j}\left(\sum_{i=1}^{n} x_i y_i - X\right) =$$

$$= \sum_{i=1}^{n} \frac{\partial x_i}{\partial \alpha_j} y_i - \sum_{i=1}^{n} \frac{\partial X}{\partial x_i}\frac{\partial x_i}{\partial \alpha_j} - \frac{\partial X}{\partial \alpha_j} = -\frac{\partial X}{\partial \alpha_j}$$

$$(j = 1, \ldots, m)$$

This completes the proof of Donkin's theorem.

We utilize Donkin's theorem to make the transition from the Lagrangian variables to the Hamiltonian variables by replacing in the theorem the function X by L, the variables $x_1, \ldots, x_n$ by, $\dot{q}_1, \ldots, \dot{q}_n$, the parameters $\alpha_1, \ldots, \alpha_m$ by $q_1, \ldots, q_n$ and t, the variables $y_1, \ldots, y_n$ by $p_1, \ldots, p_n$ and finally (taking account of the footnote 2 on page 74), the function $Y = \sum_{i=1}^{n} x_i y_i - X$ by $H = \sum_{i=1}^{n} p_i \dot{q}_i - \widehat{L}$. Then, by Donkin's theorem [the Hessian of the function L with respect to the $\dot{q}_i$ $(i = 1, \ldots, n)$ is not zero!] it follows from formulas (1) that

$$\dot{q} = \frac{\partial H}{\partial p_i}, \quad \frac{\partial L}{\partial q_i} = -\frac{\partial H}{\partial q_i} \quad (i = 1, \ldots, n) \tag{16}$$

$$\frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t} \tag{17}$$

The equations (16) and (17) are identities resulting from the relation (1) between generalized velocities and generalized momenta. But the Lagrange equations may, by virtue of (1), be written as follows:

$$\frac{dp_i}{dt} = \frac{\partial L}{\partial q_i} \quad (i = 1, \ldots, n) \tag{18}$$

---

* In computing the derivative $\frac{\partial Y}{\partial \alpha_j}$ the quantities $y_1, \ldots, y_n$ are regarded as constants.

It is these equations together with the equations (16) that lead us to the canonical equations of Hamilton:

$$\frac{dq_i}{dt}=\frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q_i} \quad (i=1, \ldots, n) \tag{19}$$

In addition to the Hamilton equations we obtained the identity (17) which will be utilized in the sequel.

From equations (19) there follows the identity

$$\frac{dH}{dt}=\sum_{i=1}^{n}\left(\frac{\partial H}{\partial q_i}\frac{dq_i}{dt}+\frac{\partial H}{\partial p_i}\frac{dp_i}{dt}\right)+\frac{\partial H}{\partial t}=\frac{\partial H}{\partial t} \tag{20}$$

We shall call a system *generalized-conservative* if the Hamiltonian function H does not explicitly depend on t, i.e. if

$$H=H(q_i, p_i)$$

In this case $\frac{\partial H}{\partial t}=0$* and, consequently, by virtue of identity (20), $\frac{dH}{dt}=0$, i.e. when the system is in motion

$$H(q_i, p_i)=\text{const}=h \tag{21}$$

where h is an arbitrary constant. We shall call the function H the *generalized total energy*, and the relation (21) which does not contain $\dot q$ or $\dot p$ and includes an arbitrary constant h, the *generalized integral of energy*.

We will explain this terminology by examining a natural system. Then L is a quadratic function [see equation (4), Sec. 11]

$$L=L_2+L_1+L_0$$

and

$$H=\sum_{i=1}^{n}p_i\widehat{\dot q}-\widehat{L}=\sum_{i=1}^{n}\widehat{\frac{\partial L}{\partial \dot q_i}\dot q_i}-\widehat{L}=$$

$$=\sum_{i=1}^{n}\widehat{\frac{\partial L_2}{\partial \dot q_i}\dot q_i}+\sum_{i=1}^{n}\widehat{\frac{\partial L_1}{\partial \dot q_i}\dot q_i}-\widehat{L}_2-\widehat{L}_1-L_0$$

But by Euler's theorem on homogeneous functions,**

$$\sum_{i=1}^{n}\frac{\partial L_2}{\partial \dot q_i}\dot q_i=2L_2, \quad \sum_{t=1}^{n}\frac{\partial L_1}{\partial \dot q_i}\dot q_i=L_1 \tag{22}$$

---

* From (17) it also follows that $\frac{\partial L}{\partial t}=0$ for a generalized-conservative system, i.e. t does not appear explicitly in the Lagrangian function L either.

** See footnote on page 49.

Therefore, for an arbitrary natural system we finally get

$$H = \widehat{L}_2 - L_0 \tag{23}$$

Let $T = T_2 + T_1 + T_0$. If the forces have the ordinary potential $\Pi = \Pi(t, q_i)$ or the generalized potential $V = V_1 + \Pi$, then $L_2 = T_2$, $L_0 = T_0 - \Pi$, and therefore

$$H = \widehat{T}_2 - T_0 + \Pi \tag{24}$$

If the system is scleronomic, then $T = T_2$, $T_0 = 0$, and so

$$H = \widehat{T} + \Pi \tag{25}$$

Thus, for a scleronomic natural system the Hamiltonian function H is the total energy[*] expressed in terms of Hamiltonian variables.

Now let us consider a conservative system, i.e. a natural scleronomic holonomic system with an ordinary force potential that is not dependent explicitly on the time.

In this case the time t does not enter into expression (25) explicitly for the total energy H and so, according to (21), we have the law of conservation of energy

$$T + \Pi = h \tag{26}$$

A conservative system is a special case of a generalized-conservative system, and in this special case the generalized integral of energy becomes an ordinary integral.

If the system is scleronomic and the forces have a generalized potential $V = V_1 + \Pi$ in which Π does not depend explicitly on the time t $\left(\frac{\partial \Pi}{\partial t} = 0\right)$, then the function H is again determined by the formula (25) and does not depend explicitly on t. For this reason, in this case too the system is a generalized-conservative system, and the energy integral (26) holds.[*]

Example. A horizontal rack is in rotation about a vertical axis, and a weight of mass m is in motion along the rack. The force acting on the weight has the potential $\Pi(r)$, where r is the distance of the weight from the axis of rotation.

Denote by φ the angle of rotation of the rack and by $I = md^2$ its moment of inertia about the vertical axis of rotation. Then r and φ are independent coordinates of the system and

$$T = \frac{1}{2} I \dot{\varphi}^2 + \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\varphi}^2) = \frac{1}{2} m [\dot{r}^2 + (r^2 + d^2) \dot{\varphi}^2]$$

From this

$$p_r = \frac{\partial T}{\partial \dot{r}} = m\dot{r}, \quad p_\varphi = \frac{\partial T}{\partial \dot{\varphi}} = m (r^2 + d^2) \dot{\varphi}$$

---

[*] Given a generalized potential $V = V_1 + \Pi$, we continue to call the quantity $T + \Pi$ (and not $T + V$) the total energy.

Since the system is conservative, it follows that

$$H = T + \Pi = \frac{1}{2m}\left(p_r^2 + \frac{p_\varphi^2}{r^2+d^2}\right) + \Pi(r)$$

The canonical equations in the given case have the form

$$\frac{dr}{dt} = \frac{p_r}{m}, \quad \frac{d\varphi}{dt} = \frac{p_\varphi}{m(r^2+d^2)},$$

$$\frac{dp_r}{dt} = \frac{rp_\varphi^2}{m(r^2+d^2)^2} - \Pi'(r), \quad \frac{dp_\varphi}{dt} = 0,$$

and the energy integral is written as

$$p_r^2 + \frac{p_\varphi^2}{r^2+d^2} + 2m\Pi(r) = h_1 \quad (h_1 = 2mh)$$

13. Routh's Equations

For the basic variables characterizing the state of a system at a time t, Routh proposed taking part of the Lagrangian variables and part of the Hamiltonian variables. The *Routh variables* are the quantities

$$t,\ q_i,\ q_\alpha,\ \dot{q}_i,\ p_\alpha \quad (i = 1, \ldots, m;\ \alpha = m+1, \ldots, n) \tag{1}$$

where m is an arbitrary fixed number less than n.

In order to pass from the Lagrangian variables to these variables, we have to express all the $\dot{q}_\alpha$ in terms of p_α $(\alpha = m+1, \ldots, n)$, utilizing the relations

$$p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha} \quad (\alpha = m+1, \ldots, n) \tag{2}$$

Suppose that the Hessian of the function L of the generalized velocities $\dot{q}_\alpha$ (the "small Hessian") is different from zero:*

$$\det\left(\frac{\partial^2 L}{\partial \dot{q}_\alpha \partial \dot{q}_\beta}\right)_{\alpha,\beta=m+1}^{n} \neq 0 \tag{3}$$

---

* In the general case this inequality does not follow from the inequality (19) on page 70 but is an auxiliary condition. Now for a natural system, the inequality (3) follows from the fact that $L_2 = \frac{1}{2}\sum_{r,s=1}^{n} a_{rs}\dot{q}_r\dot{q}_s$ is a positive definite quadratic form. In this case, the principal minor composed of coefficients of the quadratic form:

$$\det\left(\frac{\partial^2 L}{\partial \dot{q}_\alpha \partial \dot{q}_\beta}\right)_{\alpha,\beta=m+1}^{n} = \det(a_{\alpha\beta})_{\alpha,\beta=m+1}^{n}$$

must be positive.

Then, applying the Donkin theorem that was proved in the preceding section, we get a transformation that is inverse to the transformation (2), namely

$$\dot{q}_\alpha = \frac{\partial R}{\partial p_\alpha} \quad (\alpha = m+1, \ldots, n), \tag{4}$$

where $R = R(t, q_i, q_\alpha, \dot{q}_i, p_\alpha)$ is the Routh function defined by the equation

$$R = \sum_{\alpha=m+1}^{n} p_\alpha \hat{\dot{q}}_\alpha - \widehat{L}; \tag{5}$$

the sign $\frown$ signifying here that all the $\dot{q}_\alpha$ are expressed in terms of p_α.

The variables t, q_i, q_α, $\dot{q}_i$ $(i = 1, \ldots, m;\ \alpha = m+1, \ldots, n)$ are now regarded as parameters and for this reason, by virtue of the same Donkin theorem,

$$\frac{\partial R}{\partial q_i} = -\frac{\partial L}{\partial q_i}, \quad \frac{\partial R}{\partial \dot{q}_i} = -\frac{\partial L}{\partial \dot{q}_i} \quad (i = 1, \ldots, m), \tag{6}$$

$$\frac{\partial R}{\partial q_\alpha} = -\frac{\partial L}{\partial q_\alpha} \quad (\alpha = m+1, \ldots, n), \tag{7}$$

$$\frac{\partial R}{\partial t} = -\frac{\partial L}{\partial t} \tag{8}$$

Lagrange's equations of the coordinates q_i may be rewritten, in accord with the equations (6), as

$$\frac{d}{dt}\frac{\partial R}{\partial \dot{q}_i} - \frac{\partial R}{\partial q_i} = 0 \quad (i = 1, \ldots, m) \tag{9}$$

Lagrange's equations of the coordinates q_α

$$\frac{dp_\alpha}{dt} = \frac{\partial L}{\partial q_\alpha} \quad (\alpha = m+1, \ldots, n)$$

together with the formulas (4) and (7) yield

$$\frac{dq_\alpha}{dt} = \frac{\partial R}{\partial p_\alpha}, \quad \frac{dp_\alpha}{dt} = -\frac{\partial R}{\partial q_\alpha} \quad (\alpha = m+1, \ldots, n) \tag{10}$$

Equations (9) and (10)

$$\left.\begin{aligned} &\frac{d}{dt}\frac{\partial R}{\partial \dot{q}_i} - \frac{\partial R}{\partial q_i} = 0 \quad (i = 1, \ldots, m), \\ &\frac{dq_\alpha}{dt} = \frac{\partial R}{\partial p_\alpha}, \quad \frac{dp_\alpha}{dt} = -\frac{\partial R}{\partial q_\alpha} \quad (\alpha = m+1, \ldots, r) \end{aligned}\right\} \tag{11}$$

form a set of *Routh equations*. They consist of m second-order differential equations of the Lagrange type and $2(n-m)$ first-order differential equations of the Hamiltonian type; and the Routh functions in the first equations play the role of the Lagrangian function, while those in the latter equations play the role of the Hamiltonian function.*

14. Cyclic Coordinates

The coordinate q_α is called *cyclic* (ignorable) if it does not enter the Lagrangian function L explicitly, i.e. if $\frac{\partial L}{\partial q_\alpha}=0$.

When deriving the Hamilton and the Routh equations the following equalities were established:

$$-\frac{\partial L}{\partial q_\alpha}=\frac{\partial H}{\partial q_\alpha}=\frac{\partial R}{\partial q_\alpha} \tag{1}$$

From these equations it follows that the partial derivatives $\frac{\partial L}{\partial q_\alpha}$, $\frac{\partial H}{\partial q_\alpha}$ and $\frac{\partial R}{\partial q_\alpha}$ can only vanish simultaneously. For this reason, the cyclic coordinate may also be defined as a coordinate that does not appear in the Hamiltonian function H or the Routh function R explicitly. All these definitions are equivalent.

The generalized momentum that corresponds to the cyclic coordinate q_α maintains a constant value during motion. Indeed, from the canonical equations it follows that

$$\frac{dp_\alpha}{dt}=-\frac{\partial H}{\partial q_\alpha}=0$$

i.e. $p_\alpha=\text{const}=c_\alpha$.

Now suppose that we have $r=n-m$ cyclic coordinates q_α $(\alpha=m+1, \ldots, n)$. The cyclic coordinates q_α do not enter into H explicitly while the momenta p_α corresponding to these coordinates can be replaced by the constants c_α. Then**

$$H=H(t, q_i, p_i, c_\alpha) \tag{2}$$

Let us write down the canonical equations for non-cyclic coordinates:

$$\frac{dq_i}{dt}=\frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q_i} \quad (i=1, \ldots, m) \tag{3}$$

---

[*] If we wish to retain the connection between the generalized momenta and the Lagrangian function, we must take it that in the first equations of (11) the role of the Lagrangian function is played by the function $-R$, since, according to (6), $p_i=\frac{\partial L}{\partial \dot{q}_i}=\frac{\partial(-R)}{\partial \dot{q}_i}$ $(i=1, \ldots, m)$.

[**] The index i runs through the values $1, \ldots, m$, while the index α runs through $m+1, \ldots, n$.

From the structure (2) of the function H it follows that equations (3) form a system of $2m$ first-order differential equations in $2m$ unknown functions q_i, p_i $(i = 1, \ldots, m)$. Integrating this system we find

$$\left.\begin{aligned} q_i &= \varphi_i(t, c_\alpha, c_i, c_i'), \\ & \qquad\qquad (i = 1, \ldots, m), \\ p_i &= \psi_i(t, c_\alpha, c_i, c_i') \end{aligned}\right\} \tag{4}$$

where c_i, c_i' $(i = 1, \ldots, m)$ are $2m$ new arbitrary constants. By substituting expressions (4) into H we can determine the q_α from the equations

$$\frac{dq_\alpha}{dt} = \frac{\partial H}{\partial c_\alpha} \quad (\alpha = m + 1, \ldots, n) \tag{5}$$

by quadratures

$$q_\alpha = \int \frac{\partial H}{\partial c_\alpha}\, dt + c_\alpha' \quad (\alpha = m + 1, \ldots, n) \tag{6}$$

Thus, essentially the whole question reduced to integrating the system (3) whose order, $2m$, is less than the order of the initial system, $2n$, by $2r$ units, where $r = n - m$ is the number of cyclic coordinates, i.e. the presence of r cyclic coordinates made it possible to lower the order of the system by $2r$ units.

The system of equations (3) is Hamiltonian. We shall show that with the aid of the Routh equations it is possible to obtain an autonomous* system of m second-order differential equations of the Lagrangian type. Indeed, by replacing the generalized momenta p_α, which correspond to the cyclic coordinates q_α, by the constants c_α $(\alpha = m + 1, \ldots, n)$, we can write the Routh function as a function of t, q_i, $\dot{q}_i$ and c_α $(i = 1, \ldots, m;\ \alpha = m + 1, \ldots, n)$:

$$R = R(t, q_i, \dot{q}_i, c_\alpha) \tag{7}$$

Then the Routh equations for non-cyclic coordinates

$$\frac{d}{dt}\frac{\partial R}{\partial \dot{q}_i} - \frac{\partial R}{\partial q_i} = 0 \quad (i = 1, \ldots, m) \tag{8}$$

form the desired autonomous system, while the cyclic coordinates q_α are determined from the appropriate Routh equations (11) of the preceding section by means of quadratures:

$$q_\alpha = \int \frac{\partial R}{\partial c_\alpha}\, dt + c_\alpha' \tag{9}$$

---

[*] Here, by an autonomous system is meant a system of differential equations that contains no "superfluous" unknown functions which should have been determined prior to integration of the given system of equations.

Here, first all the q_i and $\dot q_i$ in $\frac{\partial R}{\partial c_\alpha}$ are replaced by the functions of $2m+1$ arguments t, c_i and c_i' $(i=1, \ldots, m)$, which functions are obtained by integrating the system (8).

Example. In the example considered at the end of Sec. 12,

$$L=T-\Pi=\frac{1}{2}m\,[\dot r^2+(r^2+d^2)\,\dot\varphi^2]-\Pi\,(r)$$

φ is a cyclic coordinate, and

$$p_\varphi=m\,(r^2+d^2)\,\dot\varphi=\text{const}=c_\varphi$$

We form the Routh function

$$R=p_\varphi\hat{\dot\varphi}-\hat L=\frac{1}{2}m\,[-\dot r^2+(r^2+d^2)\,\hat{\dot\varphi}^2]+$$

$$+\Pi\,(r)=-\frac{1}{2}m\dot r^2+\frac{1}{2m}\,\frac{p_\varphi^2}{r^2+d^2}+\Pi\,(r)=$$

$$=-\frac{1}{2}\,m\dot r^2+\frac{c_\varphi^2}{2m}\,\frac{1}{r^2+d^2}+\Pi\,(r)$$

Determining the motion reduces to integrating one second-order differential equation

$$\frac{d}{dt}\,\frac{\partial R}{\partial \dot r}-\frac{\partial R}{\partial r}=0,$$

which in expanded form reads

$$m\ddot r=\frac{c_\varphi^2}{m}\,\frac{r}{(r^2+d^2)^2}-\Pi'\,(r)$$

It will be noted that this is an equation of the relative motion of a weight along a rack, since the right-hand side involves the centrifugal force

$$\frac{c_\varphi^2}{m}\,\frac{r}{(r^2+d^2)^2}=mr\dot\varphi^2 \qquad (c_\varphi=p_\varphi)$$

Cyclic coordinates are sometimes called ignorable or kinosthenic coordinates. This name stems from the fact that in integrating the system of equations (3) or (8), we ignore the existence of the cyclic coordinates, regarding the p_α as constant parameters.

In the example we have examined, ignoring the cyclic coordinate led to ignoring the rotational motion of the rack and we obtained a differential equation for the relative motion along the rack.

The actual name "cyclic coordinate" is associated with the fact that in many problems in mechanics the angle φ which describes motion in closed trajectories (cycles) does not enter the expression for L explicitly and for this reason is a cyclic coordinate.

We note a certain analogy between a holonomic system having a cyclic coordinate and a generalized-conservative system. For the former system, $\frac{\partial H}{\partial q_\alpha} = 0$, for the latter, $\frac{\partial H}{\partial t} = 0$. For the former, we have the integral $p_\alpha = \text{const}$, for the latter, $H = \text{const}$. The roots of this analogy will be revealed later on when we consider the principal integral invariant of mechanics.

A more profound investigation of the motion of systems with cyclic coordinates will be carried out in Chapter 7.

15. The Poisson Bracket

In this section we consider some properties of integrals of a Hamiltonian system of equations of motion.

A certain function $f(t, q_i, p_i)$ is called the *integral* of the equations of motion

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad (i = 1, \ldots, n), \tag{1}$$

if for any motion of the given system this function retains a constant value c:*

$$f(t, q_i, p_i) = c \tag{2}$$

Sometimes the relation (2) itself is called the integral.

For a generalized-conservative system the function $H(q_i, p_i)$ is the integral. If the q_α are cyclic coordinates, then p_α will be the integral.

Obviously, if the functions $f_1, \ldots, f_l$ are integrals of the equations of motion, then an arbitrary function of these integrals

$$F(f_1, \ldots, f_l)$$

will also be an integral. Henceforward we shall therefore be interested only in independent integrals.

If the "complete" system of integrals consisting of $2n$ independent integrals $f_1, \ldots, f_{2n}$ (n is the number of degrees of freedom of the system) is known, then, solving the relations

$$f_k(t, q_i, p_i) = c_k \quad (k = 1, \ldots, 2n) \tag{3}$$

for q_i and p_i, we get the final equations of motion

$$q_i = \varphi_i(t, c_1, \ldots, c_{2n}), \quad p_i = \psi_i(t, c_1, \ldots, c_{2n}), \tag{4}$$

which contain $2n$ arbitrary constants $c_1, \ldots, c_{2n}$.

---

* This value of c may change when we pass from one motion of the system to another.

Thus, if $2n$ independent integrals are known, then all the motions of the system are known. If we know l independent integrals $f_1, \ldots, f_l$, where $l < 2n$, then we have only a partial idea of the motions of the system, and the greater l is, the more exhaustive our conception is. We are therefore always interested in finding the greatest possible number of independent integrals.

We shall familiarize ourselves here with the method of finding the integrals of the equations of motion proposed by Poisson and Jacobi.

Let $f(t, q_i, p_i)$ be the integral of the equations (1). Then substituting any solution of the Hamiltonian system (1) in place of q_i, p_i $(i = 1, \ldots, n)$, the function f is converted into a constant c, i.e. in accord with equations (1)

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0 \tag{5}$$

Poisson introduced a special term, the *Poisson bracket*, for the following expression composed of the partial derivatives of two arbitrary functions $\varphi(t, q_i, p_i)$ and $\psi(t, q_i, p_i)$:

$$(\varphi\psi) = \sum_{i=1}^{n} \left(\frac{\partial \varphi}{\partial q_i} \frac{\partial \psi}{\partial p_i} - \frac{\partial \varphi}{\partial p_i} \frac{\partial \psi}{\partial q_i} \right) \tag{6}$$

By means of the Poisson bracket, equation (5) — the necessary and sufficient condition for the function $f(t, q_i, p_i)$ to be the integral of the equations (1) — is written as follows:

$$\frac{\partial f}{\partial t} + (fH) = 0 \tag{7}$$

We note the following properties of the Poisson bracket:

For any functions $\varphi(t, q_i, p_i)$, $\psi(t, q_i, p_i)$, $\chi(t, q_i, p_i)$:

1°. $(\varphi\psi) = -(\psi\varphi)$;

2°. $(c\varphi\psi) = c(\varphi\psi)$ (c is a constant);

3°. $(\varphi + \psi\chi) = (\varphi\chi) + (\psi\chi)$;

4°. $((\varphi\psi)\chi) + ((\psi\chi)\varphi) + ((\chi\varphi)\psi) = 0$;

5°. $\frac{\partial}{\partial t}(\varphi\psi) = \left(\frac{\partial \varphi}{\partial t}\psi\right) + \left(\varphi\frac{\partial \psi}{\partial t}\right)$.

The identities 1°, 2°, 3°, 5° are obtained directly from the definition (6) of the Poisson bracket.

Identity 4°, which is called *Poisson's identity*, is established by means of special reasoning. Let X and Y be differential operators of the first order on the function $f(x_1, \ldots, x_m)$:

$$Xf = \sum_{k=1}^{m} X_k \frac{\partial f}{\partial x_k}, \quad Yf = \sum_{k=1}^{m} Y_k \frac{\partial f}{\partial x_k}, \tag{8}$$

where X_k, Y_k $(k=1, \ldots, m)$ are functions of the variables $x_1, \ldots, x_m$. Then the "commutator" $Z=XY-YX$ will also be a first-order operator:*

$$Zf=X(Yf)-Y(Xf)=\sum_{k=1}^{m}[X(Y_k)-Y(X_k)]\frac{\partial f}{\partial x_k} \tag{9}$$

Let us return to the Poisson bracket (φf). It may be regarded as the result of applying a linear operator Φ of the form (8) to the function f of the variables q_i, p_i $(i=1, \ldots, n)$:

$$\Phi f=(\varphi f)=\sum_{i=1}^{n}\left(-\frac{\partial\varphi}{\partial p_i}\frac{\partial f}{\partial q_i}+\frac{\partial\varphi}{\partial q_i}\cdot\frac{\partial f}{\partial p_i}\right) \tag{10}$$

This first-order differential operator is fully determined by the function φ. Analogous operators Ψ, X are determined by the functions ψ and χ.

Let us now undertake directly to establish the Poisson identity 4°. After removing the complex brackets, any term on the left-hand side of 4° will contain as a factor a second-order partial derivative of one of the functions φ, ψ, χ. But $((\varphi\psi)\chi)$ does not contain second-order partial derivatives of χ, and the sum

$$((\psi\chi)\varphi)+((\chi\varphi)\psi)=(\psi(\varphi\chi))-(\varphi(\psi\chi))=(\Psi\Phi-\Phi\Psi)_\chi$$

is a first-order differential operator relative to χ. Thus, the left-hand side of the Poisson identity does not involve second-order partial derivatives of χ, and hence, by virtue of symmetry, there are no second-order partial derivatives of φ and ψ. In other words, all the terms on the left-hand side of 4° cancel, which is what we sought to prove.

Let us now prove the basic theorem.

Jacobi-Poisson Theorem. *If f and g are integrals of equations of motion, then (fg) is also an integral of these equations.*

Proof. It is required to prove that for the function (fg) the relation (7) holds:

$$\frac{\partial(fg)}{\partial t}+((fg)H)=0 \tag{11}$$

when the same relation holds for each of the functions f, g:

$$\frac{\partial f}{\partial t}+(fH)=0,\ \frac{\partial g}{\partial t}+(gH)=0 \tag{12}$$

---

* Direct calculations show that on the right-hand side of (9) all terms involving second-order partial derivatives of the function f cancel.

Indeed, by 5°

$$\frac{\partial}{\partial t}(fg) = \left(\frac{\partial f}{\partial t}\, g\right) + \left(f\, \frac{\partial g}{\partial t}\right) = -((fH)\, g) - (f\,(gH)) = ((Hf)\, g) + ((gH)\, f)$$

Therefore, using the Poisson identity, we get

$$\frac{\partial (fg)}{\partial t} + ((fg)\, H) = ((fg)\, H) + ((gH)\, f) + ((Hf)\, g) = 0$$

thus completing the proof.

The theorem just proved yields an automatic rule that permits obtaining a third integral from the two integrals $f(t, q_i, p_i)$, $g(t, q_i, p_i)$ by means of algebraic operations and differentiation:

$$(fg) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right)$$

Taking the Poisson bracket of, for example, f and $(f\, g)$, we again get an integral, and so forth. However, it should not be forgotten that the new integral may prove to be either identically zero or a function of earlier known integrals. Thus, only by special choice of the independent integrals $f_1, \ldots, f_l$ $(l < 2n)$ can we be sure that it is possible to obtain, by means of the Poisson bracket, the integrals $f_{l+1}, \ldots, f_{2n}$ that are lacking when we want a complete system.

By way of illustration* let us consider the integrals of momenta and angular momenta for a system of free particles that is isolated from external actions:**

$$P_x = \sum p_x,\ P_y = \sum p_y,\ P_z = \sum p_z,$$
$$M_x = \sum m_x = \sum (yp_z - zp_y),\ M_y = \sum m_y = \sum (zp_x - xp_z),$$
$$M_z = \sum m_z = \sum (xp_y - yp_x),$$

where

$$p_x = m\dot{x},\ p_y = m\dot{y},\ p_z = m\dot{z}$$

The functions P_x, P_y, P_z, M_x, M_y, M_z are integrals, i.e. the "conservation integrals"

$$\begin{aligned} P_x &= c_1, & P_y &= c_2, & P_z &= c_3, \\ M_x &= c_4, & M_y &= c_5, & M_z &= c_6, \end{aligned} \tag{13}$$

hold if the system is isolated. However, in the presence of an external force field with principal vector $\boldsymbol{R}(X, Y, Z)$ and principal moment

---

* See Landau, L., Pyatigorsky, L. *Mechanics*, Moscow-Leningrad, 1940, p. 151 (in Russian).

** In this example the summation is taken over all the particles of the system.

L_0 (L_x, L_y, L_z) any of these integrals occurs if the corresponding one of the quantities X, Y, Z, L_x, L_y, L_z is equal to zero.

Form the Poisson bracket for the quantities relating to a single particle:

$$(p_x p_y)=0, \quad (p_x m_y)=-\frac{\partial p_x}{\partial p_x}\frac{\partial m_y}{\partial x}=p_z,$$

$$(m_x m_y)=\frac{\partial m_x}{\partial z}\frac{\partial m_y}{\partial p_z}-\frac{\partial m_x}{\partial p_z}\frac{\partial m_y}{\partial z}=xp_y-yp_x=m_z$$

It will be noted that the Poisson bracket, in which one quantity (p or m) refers to one particle, and the other to another particle, is always zero; for this reason,

$$\left.\begin{aligned}(P_xP_y)&=\sum(p_xp_y)=0,\\(P_xM_y)&=\sum(p_xm_y)=\sum p_z=P_z,\\(M_xM_y)&=\sum(m_xm_y)=\sum m_z=M_z\end{aligned}\right\}\tag{14'}$$

Cyclic permutation of the letters x, y, z yields similar relations:

$$\left.\begin{aligned}(P_yP_z)&=0, & (P_zP_x)&=0,\\(P_yM_z)&=P_x, & (P_zM_x)&=P_y,\\(M_yM_z)&=M_x, & (M_zM_x)&=M_y\end{aligned}\right\}\tag{14''}$$

The six laws of conservation (13) are not independent. It follows from the relations (14′) and (14″) that if we have the integrals

$$P_x=c_1, \quad M_x=c_4, \quad M_y=c_5 \tag{15}$$

then we also have the integrals

$$P_y=c_2, \quad P_z=c_3, \quad M_z=c_6 \tag{16}$$

This is all true, of course, of a potential field of force. In a nonpotential force field,

$$X=0, \quad L_x=0, \quad L_y=0 \tag{17}$$

do not yield the equations

$$Y=0, \quad Z=0, \quad L_z=0^* \tag{18}$$

---

* The conservation integrals (15) hold only if equations (17) are satisfied. Similarly, integrals (16) hold only if we have the equations (18).

CHAPTER 3

Variational Principles and Integral Invariants

16. Hamilton's Principle

We consider an arbitrary holonomic system with independent coordinates $q_1, \ldots, q_n$ and the Lagrangian function $L(t, q_i, \dot{q}_i)$.

The integral

$$W = \int_{t_0}^{t_1} L\,dt \tag{1}$$

is called the *action* (in Hamilton's sense) during a time interval (t_0, t_1) and the expression Ldt is the *elementary action**.

Since the function L is of the form $L = L(t, q_i, \dot{q}_i)$, it is necessary, in order to compute the action (1), to specify the functions $q_i = q_i(t)$ $(i = 1, \ldots, n)$ in the time interval $t_0 \leqslant t \leqslant t_1$. In other words, the action W is a functional dependent on the motion of the system.

If we arbitrarily specify the functions $q_i = q_i(t)$ $(i = 1, \ldots, n)$, then we get a certain kinematically possible motion (that is, motion admissible by the constraints). In the extended $(n+1)$-dimensional coordinate space, where the quantities q_i and the time t are the coordinates, this motion is depicted by some curve. We shall consider all possible such curves, or "paths", passing through two specified points of space $M_0(t_0, q_i^0)$ and $M_1(t_1, q_i^1)$ (see Fig. 29 for $n = 2$), i.e. all possible motions that translate the system from a given initial position (q_i^0), which it occupied at time t_0, to a given final position (q_i^1), which it occupies at time t_1. Here, from the start we fix the initial and terminal instants of time t_0 and t_1, and the initial and terminal positions of the system. The motions are otherwise arbitrary.

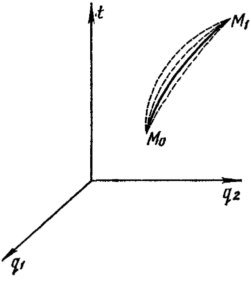

Fig. 29

---

* For natural systems, $L = T - \Pi$ has the dimensions of energy. Therefore the dimensions of action are energy × time = force × length × time.

If the system is natural and constrained, then the motions discussed here are subjected to only one restriction: during motion of the system the constraints imposed on particles of the system must not be violated. This condition is fulfilled automatically when we specify the motion in independent coordinates, assuming $q_i = q_i(t)$ $(i = 1, \ldots, n)$.

Suppose that among the paths considered there is a so-called "straight" path, i.e. one along which the system can move for a specified function L (i.e. in a given field of force). For a straight path the functions $q_i = q_i(t)$ $(i = 1, \ldots, n)$ satisfy the Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad (i = 1, \ldots, n) \tag{2}$$

All the other paths passing through the points M_0 and M_1 will be called "circuitous' paths. (In Fig. 29, the straight-line path is depicted by a solid line, the circuitous paths by dashed lines.)

We shall prove that the action W has an extremal (more precisely, stationary) value for the straight path as compared with the circuitous paths. Therein lies Hamilton's principle.*

Let us consider an arbitrary one-parameter family of paths

$$q_i = q_i(t, \alpha) \quad (t_0 \leqslant t \leqslant t_1;$$
$$-\gamma \leqslant \alpha \leqslant \gamma; \; i = 1, \ldots, n)$$

containing, at $\alpha = 0$, a given straight path; for $\alpha \neq 0$ the paths are circuitous. Let all these paths have a common origin M_0 and a common terminal point M_1:

$$q_i(t_0, \alpha) = q_i^0, \; q_i(t_1, \alpha) = q_i^1$$
$$(-\gamma \leqslant \alpha \leqslant \gamma; \; i = 1, \ldots, n)$$

The action W as computed along a path belonging to this family is a function of the parameter α·

$$W(\alpha) = \int_{t_0}^{t_1} L[t, q_i(t, \alpha), \dot{q}_i(t, \alpha)]\, dt$$

---

* This principle is given in the works of W. Hamilton published in 1834-1835. Hamilton proceeded from the assumption that the initial system was scleronomic (he proceeded from the concept of kinetic energy T as a quadratic form of the generalized velocities). This principle was formulated and substantiated for the general case of nonstationary constraints by M. V. Ostrogradsky in 1848. For this reason, the principle is sometimes called the Hamilton-Ostrogradsky principle.

We compute the variation of the action W, i.e. the differential with respect to α:

$$\delta W=\int_{t_0}^{t_1}\delta L\,dt=\int_{t_0}^{t_1}\sum_{i=1}^{n}\left(\frac{\partial L}{\partial q_i}\delta q_i+\frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i\right)dt=$$

$$=\sum_{i=1}^{n}\frac{\partial L}{\partial \dot{q}_i}\delta q_i\Big|_{t_0}^{t_1}+\int_{t_0}^{t_1}\sum_{i=1}^{n}\left(\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i\,dt=$$

$$=\int_{t_0}^{t_1}\sum_{i=1}^{n}\left(\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i\,dt \tag{3}$$

Here, we transformed the integral by means of integration by parts, making use of the commutability of the operation of variation δ and the operation of differentiation with respect to the time $\frac{d}{dt}$:

$$\delta\dot{q}_i=\delta\frac{d}{dt}q_i(t,\ \alpha)=\frac{\partial}{\partial\alpha}\frac{d}{dt}q_i(t,\ \alpha)\,\delta\alpha=$$

$$=\frac{d}{dt}\left[\frac{\partial}{\partial\alpha}q_i(t,\ \alpha)\,\delta\alpha\right]=\frac{d}{dt}\delta q_i$$

The straight path and all the circuitous paths pass through a fixed initial point and a fixed terminal point in extended coordinate space. Therefore, at $t=t_0$ and at $t=t_1$ the variations $\delta q_i=0$ and the integrated part vanishes.

From equation (3) it is seen that for the straight path, i.e. for $\alpha=0$, the expression under the sign of the transformed integral is zero by virtue of Lagrange equations. And so

$$\text{for the straight path } \delta W=0 \tag{4}$$

This is the mathematical expression of Hamilton's principle.

The converse proposition is valid as well: if for some path $\delta W=0$, then the path is straight. Indeed, because of the arbitrary nature of the variations δq_i $(i=1,\ \ldots,\ n)$ (at the ends they are zero, while at the remaining points of the path they are quite arbitrary) there follow from the condition $\delta W=0$, by virtue of the equation (3), the equations (2), i.e. the Lagrange equations for a straight path.

Since from the Hamilton principle there follow the Lagrange equations in independent coordinates (and vice versa), Hamilton's principle may be placed at the foundation of the dynamics of holonomic systems.*

---

* The statement of the Hamilton principle given above presumes (in the case of a natural system) the existence of a potential of force. The more general formulation of the principle, which embraces the case of nonpotential forces as well, will be given below [formula (9) on page 93].

The straight-line paths, i.e. the "true" motions for a given function L, may be characterized both by means of the differential equations of motion in the Lagrangian form and also by means of the variational principle of Hamilton. However, there is one fundamental difference between the differential equations of motion and the variational principles.

The differential equations of motion express a certain relation between the instant of time t, the position of the system, and the velocities and accelerations of its points at that time. If this relation holds at each point of some path, then the path is a straight line. The variational principle characterizes the entire straight path as a whole. It formulates the extremal (stationary) property of a certain functional, which property distinguishes the straight-line path from among other kinematically possible paths. The variational principle has a more surveyable and compact form and is frequently used as a foundation for new (nonclassical) domains of mechanics.

Note. The differential equations (2) are necessary and sufficient conditions for the first variation δW, where the integral W has the form (1), to be equal to zero. In the calculus of variations the equations (2) are called the *differential equations of Euler* for the variational problem

$$\delta \int_{t_0}^{t_1} L(t, q_i, \dot{q}_i)\, dt = 0$$

Lagrange's equations in independent coordinates were utilized to substantiate Hamilton's principle. But the equations themselves, in the case of a natural system, were obtained from the general equation of dynamics

$$\sum_{\nu=1}^{N} (\boldsymbol{F}_\nu - m_\nu \ddot{\boldsymbol{r}}_\nu)\, \delta \boldsymbol{r}_\nu = 0 \tag{5}$$

We shall show how Hamilton's principle can be substantiated directly by means of the general equation of dynamics (5). Then Lagrange's equations are obtainable directly from Hamilton's principle.

If in the expressions for $\boldsymbol{r}_\nu$

$$\boldsymbol{r}_\nu = \boldsymbol{r}_\nu(t, q_i) \quad (\nu = 1, \ldots, N)$$

we put in place of q_i the functions $q_i(t, \alpha)$, then $\boldsymbol{r}_\nu$ will become a complex function of t and α. We differentiate it with respect to α, i.e. we compute the variations

$$\delta \boldsymbol{r}_\nu = \sum_{i=1}^{n} \frac{\partial \boldsymbol{r}_\nu}{\partial q_i} \delta q_i \quad (\nu = 1, \ldots. N) \tag{6}$$

These formulas coincide with the formulas (8) on page 37, which determine the virtual displacements of points of a holonomic system. Thus, the variations of the radii vectors for any t are virtual displacements of the points of the system.

It is possible to convince oneself of this without resorting to formula (6) but by proceeding solely from the definition of a variation and of virtual displacement. Indeed, the variation $\delta\boldsymbol{r}_\nu = \frac{\partial\boldsymbol{r}_\nu}{\partial\alpha}\delta\alpha$ is an infinitesimal displacement of a point of the system P_ν which transfers the point of the trajectory produced for a certain fixed value of α (for a straight-line path $\alpha = 0$), to the point of an adjacent trajectory which corresponds to the parameter $\alpha + \delta\alpha$ (Fig. 30). Here, both points are taken for one and the same instant of time t, since in differentiation with respect to α the value of t is fixed. Consequently, $\delta\boldsymbol{r}_\nu$ accomplishes the transition from one possible position of the point P_ν at time t to another possible position for the same time t, i.e. $\delta\boldsymbol{r}_\nu$ is a virtual displacement of a point P_ν of the system.

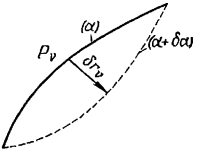

Fig. 30.

Thus, in the general equation of dynamics (5) we can consider $\delta\boldsymbol{r}_\nu$ to be a variation of the radius vector $\boldsymbol{r}_\nu$. But then it is possible to change the sequence of the operation $\frac{d}{dt}$ and the operation δ of differentiation with respect to α:

$$\frac{d}{dt}\delta\boldsymbol{r}_\nu = \frac{d}{dt}\frac{\partial}{\partial\alpha}\boldsymbol{r}_\nu(t,\ \alpha)\,\delta\alpha = \frac{\partial}{\partial\alpha}\dot{\boldsymbol{r}}_\nu(t,\ \alpha)\,\delta\alpha = \delta\dot{\boldsymbol{r}}_\nu$$

Therefore

$$\sum_{\nu=1}^{N} m_\nu\ddot{\boldsymbol{r}}_\nu\delta\boldsymbol{r}_\nu = \frac{d}{dt}\sum_{\nu=1}^{N} m_\nu\dot{\boldsymbol{r}}_\nu\delta\boldsymbol{r}_\nu - \sum_{\nu=1}^{N} m_\nu\dot{\boldsymbol{r}}_\nu\frac{d}{dt}\delta\boldsymbol{r}_\nu =$$

$$= \frac{d}{dt}\sum_{\nu=1}^{N} m_\nu\dot{\boldsymbol{r}}_\nu\delta\boldsymbol{r}_\nu - \sum_{\nu=1}^{N} m_\nu\dot{\boldsymbol{r}}_\nu\delta\dot{\boldsymbol{r}}_\nu = \frac{d}{dt}\sum_{\nu=1}^{N} m_\nu\dot{\boldsymbol{r}}_\nu\delta\boldsymbol{r}_\nu - \delta T \qquad (7)$$

where δT is a variation of the kinetic energy $T = \frac{1}{2}\sum_{\nu=1}^{N} m_\nu\dot{\boldsymbol{r}}_\nu^2$.

As before, denoting by δA the elementary work of effective forces $\boldsymbol{F}_\nu$ in virtual displacements $\delta \boldsymbol{r}_\nu$ $(\nu = 1, \ldots, N)$, we can, using the transformation (7), write down the equation (5) in the form

$$\delta T + \delta A - \frac{d}{dt} \sum_{\nu=1}^{N} m_\nu \dot{\boldsymbol{r}}_\nu \delta \boldsymbol{r}_\nu = 0$$

Integrate both sides of this equation with respect to t from $t = t_0$ to $t = t_1$:

$$\int_{t_0}^{t_1} (\delta T + \delta A)\, dt - \Big[\sum_{\nu=1}^{N} m_\nu \dot{\boldsymbol{r}}_\nu \delta \boldsymbol{r}_\nu \Big]_{t=t_0}^{t=t_1} = 0 \tag{8}$$

Here $[\quad]_{t=t_0}^{t=t_1}$ denotes the difference between the values (for $t = t_1$ and $t = t_0$) of the expression in the square brackets. But when $t = t_0$ and $t = t_1$, the radius vector $\boldsymbol{r}_\nu$ does not vary since the initial position and terminal position of the system are fixed:

$$\boldsymbol{r}_\nu (t_0, \alpha) = \boldsymbol{r}_\nu^0, \quad \boldsymbol{r}_\nu (t_1, \alpha) = \boldsymbol{r}_\nu^1 \quad (\nu = 1, \ldots, N)$$

Therefore $\delta \boldsymbol{r}_\nu = 0$ for $t = t_0$ and for $t = t_1$. The second term in (8) is zero, and this equation takes the form

$$\int_{t_0}^{t_1} (\delta T + \delta A)\, dt = 0 \tag{9}$$

We consider the case when the forces have the potential $\Pi = \Pi(t, q_i)$. Then

$$\delta A = -\delta \Pi$$

where $\delta \Pi$ is the virtual differential (variation) of the function $\Pi(t, q_i)$* and equation (9) is written

$$\int_{t_0}^{t_1} (\delta T - \delta \Pi)\, dt = \int_{t_0}^{t_1} \delta L\, dt = 0$$

whence

$$\delta W = \delta \int_{t_0}^{t_1} L\, dt = \int_{t_0}^{t_1} \delta L\, dt = 0$$

Thus, the basic equation of dynamics (5) has led us to Hamilton's principle $\delta W = 0$, and from this, as was indicated above, we immediately obtain Lagrange's equations:

---

* That is, $\delta \Pi = \sum_{i=1}^{n} \frac{\partial \Pi}{\partial q_i} \delta q_i$.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad (i = 1, \ldots, n)$$

In order to obtain Lagrange's equations, in the case of the nonpotential forces Q_i one should proceed from the equation (9) [in place of (4)]. Applying formula (3) to the integral $\int_{t_0}^{t_1} \delta T\, dt$, i.e. replacing the function L by the function $T(t, q_i, \dot{q}_i)$, and using the expression for the elementary work of the effective forces $\delta A = \sum_{i=1}^{n} Q_i \delta q_i$, we find

$$\int_{t_0}^{t_1} \sum_{i=1}^{n} \left(Q_i + \frac{\partial T}{\partial q_i} - \frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i}\right) \delta q_i\, dt = 0 \tag{10}$$

From this it follows that since the quantities δq_i $(i = 1, \ldots, n)$ are arbitrary, the expressions in the brackets under the integral sign (10) must be zero, i.e the Lagrange equations must hold for the straight-line path:

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = Q_i \quad (i = 1, \ldots, n) \tag{11}$$

Let us find out whether the value of the action for a straight-line path is the least when compared with circuitous paths.

By way of illustration let us consider the motion of a particle constrained to move on a sphere in the absence of a field of force (inertial motion on a sphere). Let the mass of the particle $m = 1$. In spherical coordinates (Fig. 31)

$$L = \frac{v^2}{2} = \frac{R^2}{2}(\dot{\varphi}^2 + \sin^2\varphi\, \dot{\psi}^2)$$

For the straight-line part

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\varphi}} - \frac{\partial L}{\partial \varphi} = 0 \quad \text{and} \quad \frac{\partial L}{\partial \dot{\psi}} = \text{const}$$

(where ψ is a cyclic coordinate), i.e.

$$\ddot{\varphi} - \sin\varphi \cos\varphi\, \dot{\psi}^2 = 0, \quad \sin^2\varphi\, \dot{\psi} = \sin^2\varphi_0\, \dot{\psi}_0$$

Without loss of generality, we may say that the initial velocity v_0 for a straight path is directed along a meridian ($\psi = \text{const}$), i.e. $\dot{\psi}_0 = 0$; then

$$\dot{\psi} = 0, \quad \dot{\varphi} = \text{const} \quad \text{and} \quad v^2 = R^2\dot{\varphi}^2 = \text{const}$$

Thus, the straight-line path amounts to uniform motion along an arc of a great circle. Here,

$$W_{cir}-W_{st}=\frac{1}{2}\int_{t_0}^{t_1}(v^2-v_0^2)\,dt=$$

$$=v_0\int_{t_0}^{t_1}(v-v_0)\,dt+\frac{1}{2}\int_{t_0}^{t_1}(v-v_0)^2\,dt\geqslant v_0\int_{t_0}^{t_1}(v-v_0)\,dt=v_0(\sigma_{cir}-\sigma_{st})$$

The length of the arc of a great circle σ_{st} is less than the length σ_{cir} of any other curve on the sphere connecting the same two points M_0 and M_1. For this reason,

$$W_{st}<W_{cir}$$

However, this is valid only when $\sigma_{st}=\smile M_0M_1<\pi R$. If $\sigma_{st}>\pi R$, then W_{st} will no longer be always less than W_{cir} and the least value of the action W will be attained on an auxiliary arc of a great circle, which in this case will be the shortest distance between M_0 and M_1. If we move the point M_1 along an arc of a great circle, increasing the arc, then the critical point M_0^* (prior to this point, W_{st} will be a minimum, and after M_1 passes through M_0^* it will no longer be such) is a point diametrically opposite the point M_0.

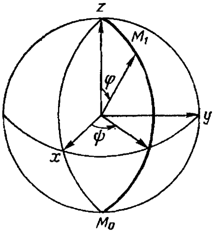

Fig. 31

The situation is similar in the general case. It may be proved that if the point M_1 is chosen sufficiently close to M_0, then one straight-line path passes through M_0 and M_1. But if the separation between M_1 and M_0 is sufficient, then two straight-line paths can pass through M_0 and M_1 or even a pencil of straight-line paths. Such a position M_0^* of the point M_1 is called the *conjugate kinetic focus* of M_0. It has been established that an action directed along the straight-line path M_0M_1 is of least value compared with those along circuitous pathways if on the arc M_0M_1 there is no kinetic focus M_0^* conjugate to M_0.

17. Second Form of Hamilton's Principle

Let us examine yet another form of Hamilton's variational principle. In place of an $(n+1)$-dimensional extended coordinate space we consider a $(2n+1)$-dimensional extended phase space in which the quantities q_i, p_i $(i = 1, \ldots, n)$ and t are the coordinates of a point. In this space we consider a straight-line path passing through two points $B_0\ (q_i^0, p_i^0, t_0)$ and $B_1\ (q_i^1, p_i^1, t_1)$ and also all other possible curves connecting these points (circuitous pathways; see Fig. 32, $n = 1$).

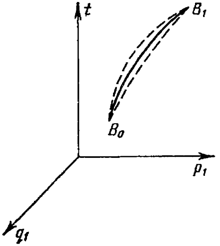

Fig. 32

The functions $q_i(t)$ and $p_i(t)$ $(i = 1, \ldots, n)$, which specify the straight-line path, satisfy Hamilton's equations

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}$$
$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad (i = 1, \ldots, n) \tag{1}$$

We introduce a function of $4n+1$ independent variables $L^*(t, q_i, p_i, \dot{q}_i, \dot{p}_i)$ defined by the equation*

$$L^* = \sum_{i=1}^{n} p_i \dot{q}_i - H(t, q_i, p_i) \tag{2}$$

---

* The $\dot{p}_i$ actually are not involved in the right-hand side of (2). Therefore, the function L^* is, in this case, independent of these quantities, and $\frac{\partial L^*}{\partial \dot{p}_i} = 0$ $(i = 1, \ldots, n)$.

With the aid of this function, Hamilton's canonical equations (1) can be written, as is readily seen, in the Lagrangian form:

$$\frac{d}{dt}\frac{\partial L^*}{\partial \dot{q}_i}-\frac{\partial L^*}{\partial q_i}=0,\quad \frac{d}{dt}\frac{\partial L^*}{\partial \dot{p}_i}-\frac{\partial L^*}{\partial p_i}=0 \quad (i=1,\ \ldots,\ n) \tag{3}$$

Since a straight-line path is characterized by equations (3) of the Lagrangian type, the straight-line path in extended phase space stands out among the circuitous pathways, as was earlier established, by the fact that for it the integral

$$\int_{t_0}^{t_1} L^*(t,\ q_i,\ p_i,\ \dot{q}_i,\ \dot{p}_i)\,dt \tag{4}$$

has the stationary value

$$\delta\int_{t_0}^{t_1} L^*\,dt=\delta\int_{t_0}^{t_1}\Big(\sum_{i=1}^{n} p_i\dot{q}_i-H\Big)\,dt=0 \tag{5}$$

At first glance it might appear that the second form (5) of Hamilton's principle does not differ in any way from the first, $\delta W = 0$ since, according to formula (8) on page 73, the expression for L^* coincides with the function L. This is not always so, however; it is true only for motions of a system, i.e. for paths such that $q_i = q_i(t)$, $p_i = p_i(t)$ $(i = 1, \ldots, n)$, for which the functions $q_i(t)$ and $p_i(t)$ are connected by the relations

$$p_i=\frac{\partial L}{\partial \dot{q}_i} \quad (i=1,\ \ldots,\ n) \tag{6}$$

However, in the second form of Hamilton's principle (in contrast to the first!), arbitrary curves of $(2n + 1)$-dimensional extended phase space passing through the points B_0 and B_1 are admitted for comparison as circuitous pathways. These paths may not satisfy relations (6) and for this reason, in the general case, for them $L^* \neq L$. But if in formula (5) we confine ourselves solely to those circuitous paths for which (6) holds, then the second form of Hamilton's principle passes into the first form, $\delta W = 0$.

Note also that unlike the points M_0 and M_1 in the first form of Hamilton's principle, points B_0 and B_1 cannot be chosen in arbitrary fashion because it is impossible in the general case to draw a straight-line path through two arbitrary points of extended phase space. The points B_0 and B_1 are chosen on the straight-line path for which Hamilton's principle is formulated.

18. The Basic Integral Invariant of Mechanics (Poincaré-Cartan Integral Invariant)

Let us derive a formula for the variation of action δW in the general case when the initial and terminal instants of time, just like the initial and terminal coordinates, are not fixed but are functions of a parameter α:

$$\left.\begin{array}{ll} t_0 = t_0(\alpha), & q_i^0 = q_i^0(\alpha) \\ t_1 = t_1(\alpha), & q_i^1 = q_i^1(\alpha) \end{array} \quad (i = 1, \ldots, n) \right\} \tag{1}$$

Here, differentiating the integral $W = \int_{t_0}^{t_1} L\,dt$ with respect to the parameter α and integrating by parts we find

$$\delta W = \delta \int_{t_0(\alpha)}^{t_1(\alpha)} L\,dt = L_1\delta t_1 - L_0\delta t_0 + \int_{t_0}^{t_1} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i\right) dt = L_1\delta t_1 + \sum_{i=1}^{n} p_i^1 [\delta q_i]_{t=t_1} - L_0\delta t_0 - \sum_{i=1}^{n} p_i^0 [\delta q_i]_{t=t_0} + \int_{t_0}^{t_1} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right) \delta q_i\,dt \tag{2}$$

$$[\delta q_i]_{t=t_\lambda} = \left[\frac{\partial}{\partial \alpha} q_i(t, \alpha)\right]_{t=t_\lambda} \delta\alpha \quad (i = 1, \ldots, n;\ \lambda = 0, 1) \tag{3}$$

On the other hand, for the variations of the terminal coordinates $q_i^1 = q_i^1[t_1(\alpha), \alpha]$ we have the formulas

$$\delta q_i^1 = \dot{q}_i^1 \delta t_1 + \left[\frac{\partial q_i(t, \alpha)}{\partial \alpha}\right]_{t=t_1} \delta\alpha$$

or

$$\delta q_i^1 = [\delta q_i]_{t=t_1} + \dot{q}_i^1 \delta t_1 \quad (i = 1, \ldots, n)$$

From this

$$[\delta q_i]_{t=t_1} = \delta q_i^1 - \dot{q}_i^1 \delta t_1 \quad (i = 1, \ldots, n) \tag{4}$$

In similar fashion

$$[\delta q_i]_{t=t_0} = \delta q_i^0 - \dot{q}_i^0 \delta t_0 \quad (i = 1, \ldots, n) \tag{5}$$

Substituting the expressions (4) and (5) for $[\delta q_i]_{t=t_1}$ and $[\delta q_i]_{t=t_0}$ into expression (2) for δW, expressing as usual $\dot{q}_i$ in terms of p_i, and noting that

$$\sum_{i=1}^{n} p_i \widehat{\dot{q}}_i - \widehat{L} = H$$

we obtain the following formula for the variation of action δW in the general case:

$$\delta W = \Big[\sum_{i=1}^{n} p_i \delta q_i - H\delta t\Big]_0^1 + \int_{t_0}^{t_1} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i\, dt \qquad (6)$$

where

$$\Big[\sum_{i=1}^{n} p_i \delta q_i - H\delta t\Big]_0^1 = \sum_{i=1}^{n} p_i^1 \delta q_i^1 - H_1 \delta t_1 - \sum_{i=1}^{n} p_i^0 \delta q_i^0 + H_0 \delta t_0$$

In the special case when for any α the appropriate path is a straight line, i.e. when $q_i = q_i(t, \alpha)$ $(i = 1, \ldots, n)$ is a family of straight-line paths, the integral on the right of (6) is equal to zero for any α, and the formula for the variation of the action takes on the following simple form:

$$\delta W = \Big[\sum_{i=1}^{n} p_i \delta q_i - H\delta t\Big]_0^1 \qquad (7)$$

In place of the $(n + 1)$-dimensional extended coordinate space we take the $(2n + 1)$-dimensional extended phase space in which the quantities q_i, p_i $(i = 1, \ldots, n)$ and t will be the coordinates of the point. In this space we take an arbitrary closed curve C_0 given by the equations

$$q_i = q_i^0(\alpha), \quad p_i = p_i^0(\alpha), \quad t = t_0(\alpha) \qquad (8)$$

$$(i = 1, \ldots, n;\ 0 \leqslant \alpha \leqslant l)$$

Here, we have one and the same point of the curve C_0 for $\alpha = 0$ and $\alpha = l$. Taking each point on the curve C_0 as the initial one we draw the appropriate straight-line path. Such a path is uniquely determined (after specifying the initial point) by a system of Hamilton's canonical equations. We obtain a closed tube of straight trajectories (see Fig. 33, $n = 1$)

$$q_i = q_i(t, \alpha), \; p_i = p_i(t, \alpha) \quad (i = 1, \ldots, n; \; 0 \leqslant \alpha \leqslant l), \tag{9}$$

where

$$q_i(t, 0) \equiv q_i(t, l), \; p_i(t, 0) \equiv p_i(t, l)$$
$$(i = 1, \ldots, n)$$

On this tube we arbitrarily choose a second closed curve C_1 around the tube that has only one common point with each generatrix.

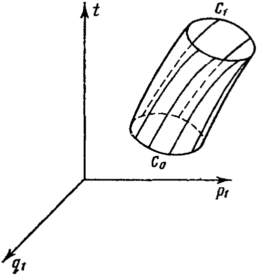

Fig. 33

The equations of curve C_1 may be written in the form*

$$q_i = q_i^1(\alpha), \; p_i = p_i^1(\alpha), \; t = t_1(\alpha) \tag{10}$$

Let us examine the action W along the generatrix of the tube from the curve C_0 to the curve C_1:

$$W = \int_{t_0(\alpha)}^{t_1(\alpha)} L\,dt$$

Then for any α, by formula (7),

$$\delta W = W'(\alpha)\,\delta\alpha = [\sum_{i=1}^{n} p_i \delta q_i - H \delta t]_0^1$$

---

* To each value of α of the interval $0 \leqslant \alpha \leqslant l$ there corresponds a definite "generatrix" of the tube (straight-line path) and on this generatrix there is only one point belonging to the curve C_1. Therefore, to each value of α there corresponds only one point of the curve C_1, i.e. the coordinates of the point of the curve C_1 are functions of the parameter α.

Integrating this equation termwise with respect to α from $\alpha=0$ to $\alpha=l$, we get

$$0=W(l)-W(0)=\int_0^l\left[\sum_{i=1}^n p_i\delta q_i-H\delta t\right]_0^1=$$

$$=\int_0^l\left[\sum_{i=1}^n p_i^1\delta q_i^1-H_1\delta t_1\right]-\int_0^l\left[\sum_{i=1}^n p_i^0\delta q_i^0-H_0\delta t_0\right]=$$

$$=\oint_{C_1}\left[\sum_{i=1}^n p_i\delta q_i-H\delta t\right]-\oint_{C_0}\left[\sum_{i=1}^n p_i\delta q_i-H\delta t\right],$$

i.e.

$$\oint_{C_0}\left[\sum_{i=1}^n p_i\delta q_i-H\delta t\right]=\oint_{C_1}\left[\sum_{i=1}^n p_i\delta q_i-H\delta t\right] \tag{11}$$

It is thus established that the line integral

$$I=\oint\left[\sum_{i=1}^n p_i\delta q_i-H\delta t\right] \tag{12}$$

taken along an arbitrary closed contour does not change its value in the case of an arbitrary displacement (with deformation) of the contour along a tube of straight-line trajectories, i.e. it is an *integral invariant*. We shall call the integral I the *Poincaré-Cartan integral invariant.*

We now prove the converse proposition. Let it be known that the straight-line paths are defined by the following system of first-order differential equations:

$$\frac{dq_i}{dt}=Q_i(t,\ q_j,\ p_j),\quad \frac{dp_i}{dt}=P_i(t,\ q_j,\ p_j)$$

$$(i=1,\ \ldots,\ n) \tag{13}$$

This supposition is natural in that the motion of a system must be defined uniquely on the basis of the initially given q_i^0, p_i^0 $(i=1,\ \ldots\ \ldots,\ n)$. Besides, let it be given that the Poincaré-Cartan integral (12) is an integral invariant with respect to the straight paths defined by the set of equations (13), i.e. for any tube of these paths the Poincaré-Cartan integral computed round the closed contour embracing the tube does not change its magnitude when the points of the contour are arbitrarily displaced along the generatrices of the tube. We will now prove that the following relations hold between the function H and the functions Q_i, P_i:

$$Q_i = \frac{\partial H}{\partial p_i}, \quad P_i = -\frac{\partial H}{\partial q_i} \quad (i = 1, \ldots, n) \tag{14}$$

That is to say, we will prove that equations (13) are Hamilton's canonical equations with the function H that enters into the expression under the sign of the integral I.

To proceed with the proof we introduce an auxiliary variable (parameter) μ, supplementing the system (13) with yet another equation:

$$\frac{dq_1}{Q_1} = \ldots = \frac{dq_n}{Q_n} = \frac{dp_1}{P_1} = \ldots = \frac{dp_n}{P_n} = \frac{dt}{1} = \pi \, d\mu \tag{15}$$

Here, $\pi = \pi(t, q_i, p_i)$ is an arbitrary function of a point in the extended phase space. Integrating the system (15), we find the expressions for q_i, p_i and t in the form of functions of the variable μ and of the arbitrary initially given values, q_i^0, p_i^0 $(i = 1, \ldots, n)$, and t_0 (for $\mu = 0$):

$$\left.\begin{aligned} q_i &= \varphi_i(\mu;\, q_j^0,\, p_j^0,\, t_0), \\ p_i &= \psi_i(\mu;\, q_j^0,\, p_j^0,\, t_0), \ (i = 1, \ldots, n) \\ t &= \chi(\mu;\, q_i^0,\, p_i^0,\, t_0) \end{aligned}\right\} . \tag{16}$$

We have thus obtained parametric equations for the family of all straight-line paths. Since we only need those straight-line paths that form the given tube, we have to choose the initial point $M_0(q_j^0, p_j^0, t_0)$ on the curve C_0, i.e. we have to put

$$q_j^0(\alpha), \quad p_j^0(\alpha), \quad t_0(\alpha)$$

into equations (16) in place of q_j^0, p_j^0 and t_0. Having done that we will find the parametric equations for the straight-line paths that form the given tube:

$$\begin{gathered} q_i = q_i(\mu, \alpha), \quad p_i = p_i(\mu, \alpha), \\ t = t(\mu, \alpha) \quad (i = 1, \ldots, n;\ 0 \leqslant \alpha \leqslant l) \end{gathered} \tag{17}$$

Here, the value of α isolates a definite straight-line path (the "generatrix" of the tube), while the value of the parameter μ fixes a definite point on this path.

Assuming $\mu = \text{const}$, we obtain a point on every generatrix, and on the tube we get a closed curve. We take it that in the integral (12), q_i, p_i and t are replaced by their expressions (17). Then the integral I will be a function of the parameter μ and for every fixed value of μ, it will be a line integral along the corresponding closed curve $\mu = \text{const}$.

By virtue of invariance,

$$dI=0,$$

where the letter d denotes differentiation with respect to the parameter μ. Differentiating under the sign of the integral, we find

$$0=\oint\Big[\sum_{i=1}^{n}(dp_i\delta q_i+p_i\,d\delta q_i)-dH\delta t-H\,d\delta t\Big]$$

Writing $\delta\,dq_i$ and $\delta\,dt$ in place of $d\delta q_i$ and $d\delta t$, and integrating by parts round the closed contour, we get*

$$0=\oint\Big[\sum_{i=1}^{n}(dp_i\delta q_i-\delta p_i\,dq_i)-dH\delta t+\delta H\,dt\Big]=$$

$$=\oint\Big\{\sum_{i=1}^{n}\Big[\Big(dp_i+\frac{\partial H}{\partial q_i}dt\Big)\delta q_i+\Big(-dq_i+\frac{\partial H}{\partial p_i}dt\Big)\delta p_i\Big]+$$

$$+\Big(-dH+\frac{\partial H}{\partial t}dt\Big)\delta t\Big\},$$

or, by virtue of (15), dividing termwise by $d\mu=\frac{dt}{\pi}$,

$$0=\oint\Big\{\sum_{i=1}^{n}\Big[\Big(P_i+\frac{\partial H}{\partial q_i}\Big)\delta q_i+\Big(-Q_i+\frac{\partial H}{\partial p_i}\Big)\delta p_i\Big]+$$

$$+\Big(-\frac{dH}{dt}+\frac{\partial H}{\partial t}\Big)\delta t\Big\}\pi \qquad (18)$$

The expression under the integral sign must be a total differential with an arbitrary factor π; but this is only possible when the expression in the braces is zero. Equating this expression to zero, we get

$$P_i=-\frac{\partial H}{\partial q_i},\quad Q_i=\frac{\partial H}{\partial p_i}\qquad (i=1,\ \ldots,\ n),$$

which completes the proof.**

---

* The operations d and δ may be interchanged since they represent differentiation with respect to different independent variables μ and α. Further, when integrating by parts, the integrated part is lost (is equal to zero) since the terminal point of the integration path coincides with the initial point. Thus, for any two functions u and v, $\oint u\delta v=-\oint v\delta u$ when integrating over a closed contour.

** We also obtain the identity $\frac{dH}{dt}=\frac{\partial H}{\partial t}$, which is a consequence of Hamilton's canonical equations [see equation (20) on page 76].

From what has been proved it follows that the invariance of the Poincaré-Cartan integral may be placed at the foundation of mechanics since from this invariance it follows that the motion of a system obeys Hamilton's canonical equations.

Note. In the proof we introduced an auxiliary variable μ and used the circumstance that the integral I does not change its value in passing from one curve of the family $\mu = \text{const}$ to another curve of the same family. Because the function $\pi(t, q_i, p_i)$ is arbitrary, the family of curves $\mu = \text{const}$ is actually an arbitrary family of nonintersecting closed curves embracing the given tube of straight-line paths. If we had not introduced the parameter μ, and had taken the time t for the parameter, then, by the same reasoning, we would have only partially made use of the invariance of the integral I (only for the curves of the simultaneous states $t = \text{const}$) and we could not have arrived at the desired result.

Now let us take a more detailed look at the structure of the Poincaré-Cartan integral.

In the Poincaré-Cartan integral (12) the time t enters as the coordinate q_i, and the role of the corresponding momentum is played by the quantity $-H$, i.e. the energy taken with opposite sign. This is a far-reaching analogy.

We change the variables in the integral I by introducing a new variable z connected with the old variables by the relation

$$z = -H(t, q_i, p_i) \tag{19}$$

Using this relation let us express p_1:

$$p_1 = -K(t, q_1, \ldots, q_n, z, p_2, \ldots, p_n) \tag{20}$$

Then the basic integral I in the new variables will read

$$I = \oint z\delta t + p_2\delta q_2 + \ldots + p_n\delta q_n - K\delta q_1 \tag{21}$$

Thus, in the new variables the integral I has the aspect of the Poincaré-Cartan integral, but the role of the time is now played by the variable q_1 and in place of the earlier energy H we have the momentum p_1 taken with reversed sign, i.e. K. Thus, by what has been proved, the motion of a system in the new variables should be described by the following Hamiltonian system of differential equations:

$$\left.\begin{aligned} &\frac{dt}{dq_1} = \frac{\partial K}{\partial z}, \quad \frac{dz}{dq_1} = -\frac{\partial K}{\partial t}, \\ &\frac{dq_j}{dq_1} = \frac{\partial K}{\partial p_j}, \quad \frac{dp_j}{dq_1} = -\frac{\partial K}{\partial q_j} \quad (j = 2, \ldots, n) \end{aligned}\right\} \tag{22}$$

Where, q_1 is the independent variable.

We can illustrate this in the case of a linear oscillator for which

$$H=\frac{p^2}{2m}+\frac{cq^2}{2}.$$

Form the canonical equations taking q for the independent variable. To do this, put

$$z=-\left(\frac{p^2}{2m}+\frac{cq^2}{2}\right)$$

whence

$$p=\sqrt{m}\sqrt{-2z-cq^2}$$

Thus, we have

$$K=-\sqrt{m}\sqrt{-2z-cq^2}$$

The appropriate canonical equations (22) will then read

$$\frac{dt}{dq}=\sqrt{\frac{m}{c}}\frac{1}{\sqrt{-2\frac{z}{c}-q^2}}, \quad \frac{dz}{dq}=0$$

From the second equation, $z=\text{const}=-h$. We find t by means of a *quadrature*:

$$\omega t=\int\frac{dq}{\sqrt{\frac{2h}{c}-q^2}}-\alpha$$

where $\omega=\sqrt{\frac{c}{m}}$, and α is an arbitrary constant, or

$$\omega t+\alpha=\arcsin\sqrt{\frac{c}{2h}}\,q$$

i.e.

$$q=A\sin(\omega t+\alpha)\ \left(A=\sqrt{\frac{2h}{c}}\right)$$

19. A Hydrodynamical Interpretation of the Basic Integral Invariant. The Theorems of Thomson and Helmholtz on Circulation and Vortices

For a concrete interpretation of the concept of an integral invariant let us consider the motion of an ideal fluid under the action of external forces with a potential $\Pi(t, x, y, z)$. As we know from hydrodynamics the equation of motion of a particle of this fluid has the form

$$a=-\operatorname{grad}\Pi-\frac{1}{\rho}\operatorname{grad}p, \tag{1}$$

where a is the acceleration of the particle, ρ and p are its density and pressure, while the potential Π is referred to unit mass.

We take it that ρ and p are connected by the functional relation $\rho = f(p)$ (in particular this occurs in an isothermal development of the process). Then, putting

$$\tilde{\Pi} = \Pi + \int \frac{dp}{\rho}$$

we can write equation (1) as

$$a = -\operatorname{grad} \tilde{\Pi} \tag{1'}$$

This equation shows that the motion of a particle of the fluid is identical to the motion of a particle with mass $m = 1$ in a potential field $\tilde{\Pi} = \tilde{\Pi}(t, x, y, z)$. Therefore, the integral invariant for the motion of fluid particles will be the Poincaré-Cartan integral, which in the given case is of the form

$$I = \oint u\delta x + v\delta y + w\delta z - E\delta t \tag{2}$$

where u, v, and w are the components of velocity of the particle (in the given case—for $m = 1$—they are the momenta p_i), and E is the energy defined by the formula

$$E = \frac{1}{2}(u^2 + v^2 + w^2) + \tilde{\Pi}(t, x, y, z) \tag{3}$$

Thus, the integral (2) taken round an arbitrarily closed contour in the seven-dimensional space (t, x, y, z, u, v, w) does not change its magnitude in an arbitrary displacement of points of the contour in accordance with the motion of the fluid. This motion obeys the differential equations, which, by virtue of the formula (1′), have the form

$$\left.\begin{array}{c} \dfrac{dx}{dt} = u, \quad \dfrac{dy}{dt} = v, \quad \dfrac{dz}{dt} = w, \\ \dfrac{du}{dt} = -\dfrac{\partial \tilde{\Pi}}{\partial x}, \quad \dfrac{dv}{dt} = -\dfrac{\partial \tilde{\Pi}}{\partial y}, \quad \dfrac{dw}{dt} = -\dfrac{\partial \tilde{\Pi}}{\partial z} \end{array}\right\} \tag{4}$$

For the given special case, equations (4) are Hamilton's canonical equations.

If the specific motion of a fluid is given for which the field of velocities is known, i.e. if we know the functions $u(t, x, y, z)$, $v(t, x, y, z)$, and $w(t, x, y, z)$, then the integral (2) may be regarded as an integral in extended coordinate space, i.e. in the four-dimensional space t, x, y, z. The value of this integral does not change if the points of the path of integration are displaced arbitrarily along the paths of motion of the particles; i.e. the integral

$$\oint_C u(t, x, y, z)\,dx + v(t, x, y, z)\,dy + \\ + w(t, x, y, z)\,dz - E(t, x, y, z)\,dt \tag{5}$$

is an integral invariant in extended coordinate space for the motion of a fluid with a given field of velocities.

If the path of integration consists of simultaneous states ($t = \text{const}$), then the integral (2) takes the form

$$\oint_C u\delta x + v\delta y + w\delta z \tag{6}$$

In hydrodynamics this integral is called the *circulation of the velocity* along the contour C. Along with this we obtained **Thomson's theorem** on the conservation of the circulation of velocity: *the magnitude of the circulation* (6) *does not change if the fluid particles forming the contour at time t_1 are transferred to positions that they occupy at any other arbitrary instant of time t_2.*

If the fluid particles at some time form a line, then these same particles will form another line at another instant of time. We shall speak of a "fluid line" that moves and is deformed with time. In similar fashion we define the concept of a "fluid surface".

The theorem on the conservation of circulation states that *to every closed fluid line there corresponds a definite circulation.*

Note that by the Stokes formula the integral (6) is written in the form of an integral over the surface S bounded by the contour C:

$$\iint_S \xi\delta y\delta z + \eta\delta z\delta x + \zeta\delta x\delta y, \tag{7}$$

where

$$\xi = \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}, \quad \eta = \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}, \\ \zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \tag{8}$$

are the components of some vector $\mathbf{\Omega}$ called the *curl* (*rotor*) *of the velocity* or simply the *curl*. Integral (7) is ordinarily given in the form

$$\iint_S \Omega_n\,dS, \tag{7'}$$

where Ω_n is the projection of the vector $\mathbf{\Omega}$ on the normal to the surface, and dS is an element of the surface S. From this it is seen that the integral (7) gives the magnitude of the vortex flow through the surface. The theorem on the conservation of the circulation of velocity passes into the theorem on the conservation of vortex flow: to every bounded fluid surface there corresponds a definite magnitude of vortex flow through this surface.*

The motion of a fluid with a given field of velocities is determined by the differential equations

$$\begin{aligned}&\frac{dx}{dt}=u(t,\ x,\ y,\ z),\\ &\frac{dy}{dt}=v(t,\ x,\ y,\ z),\quad \frac{dz}{dt}=w(t,\ x,\ y,\ z)\end{aligned}\tag{9}$$

With respect to the integral curves of the system (9), the integral (5) is an integral invariant.

We now pose the question of what other systems of differential equations of the type

$$\frac{dx}{P(t,\ x,\ y,\ z)}=\frac{dy}{Q(t,\ x,\ y,\ z)}=\frac{dz}{R(t,\ x,\ y,\ z)}=\frac{dt}{U(t,\ x,\ y,\ z)}\tag{10}$$

besides the system (9), possess this property; in other words, for what other systems is the integral (5) an integral invariant?

To answer this question, let us introduce a parameter μ for the trajectories of the system (10) and, as was done in the preceding section, let us equate the relations (10) to the product $\pi(t,\ x,\ y,\ z)\,d\mu$, where π is an arbitrary function. Let us consider a tube of integral curves of the system (10) and a closed contour C around this tube, for which contour $\mu=\text{const}=c$. Note that the value of the integral (5) along the contour C does not depend on the quantity $\mu=c$.

Denoting by d the differentiation with respect to μ and reasoning as on page 103 we get [using formulas (8)]

$$\begin{aligned}0&=d\oint u\delta x+v\delta y+w\delta z-E\delta t=\oint du\delta x+\partial v\delta y+\\ &\quad+dw\delta z-dE\delta t-\delta u\,dx-\delta v\,dy-\delta w\,dz+\delta E\,dt=\\ &=\oint\Big[-\zeta\,dy+\eta\,dz+\Big(\frac{\partial u}{\partial t}+\frac{\partial E}{\partial x}\Big)dt\Big]\delta x+[*]\,\delta y+[*]\,\delta z+\\ &\quad+\Big[-dE-\frac{\partial u}{\partial t}\,dx-\frac{\partial v}{\partial t}\,dy-\frac{\partial w}{\partial t}\,dz+\frac{\partial E}{\partial t}\,dt\Big]\delta t\end{aligned}$$

where the coefficients of δy and δz are obtained from the coefficients of δx by a cyclic permutation.

---

* From this it follows in particular that vortices in the fluid volume of an ideal fluid can neither vanish nor appear [if of course the forces have a potential and the relation $\rho=f(p)$ holds].

Replace dx, dy, dz, dt by the denominators of the fractions in (10) multiplied by $\pi(t, x, y, z)\,dt$. Since for any choice of the function $\pi(t, x, y, z)$ the expression under the integral sign must be a total differential, it must be identically zero. And so the expressions in square brackets are equal to zero (after the differentials dx, dy, dz, dt in them are replaced by the proportional quantities P, Q, R, U); i.e. we have the equations

$$\left.\begin{aligned} &\eta R-\zeta Q+\left(\frac{\partial u}{\partial t}+\frac{\partial E}{\partial x}\right)U=0\\ &\zeta P-\xi R+\left(\frac{\partial v}{\partial t}+\frac{\partial E}{\partial y}\right)U=0\\ &\xi Q-\eta P+\left(\frac{\partial w}{\partial t}+\frac{\partial E}{\partial z}\right)U=0 \end{aligned}\right\} \qquad (11)$$

and

$$\left(\frac{\partial u}{\partial t}+\frac{\partial E}{\partial x}\right)P+\left(\frac{\partial v}{\partial t}+\frac{\partial E}{\partial y}\right)Q+\left(\frac{\partial w}{\partial t}+\frac{\partial E}{\partial z}\right)R=0 \qquad (12)$$

The relation (12) is a consequence of the equations (11) if $U \neq 0$, and of the equations (11) and (1′) if $U = 0$.

The equations (11) together with equations (10) determine all the differential systems with respect to which the integral (5) is an integral invariant. Among these systems we shall seek those for which $U = 0$, i.e. $dt = 0$. Then from (11) we get

$$\frac{P}{\xi}=\frac{Q}{\eta}=\frac{P}{\zeta},$$

and the system (10) takes the form

$$\frac{dx}{\xi}=\frac{dy}{\eta}=\frac{dz}{\zeta}=\frac{dt}{0} \qquad (13)$$

The integral curves of the system (13) are called *vortex lines*.

Thus, *the system of differential equations of vortex lines is the only system with* $dt = 0$, *with respect to which the integral* (5) *is an integral invariant.**

From this important proposition follows an important corollary.

In the space (x, y, z, t) take an arbitrary tube of vortex lines and two contours C_1 and C_2 bounding it (Fig. 34). By virtue of the invariance of the integral (5) with respect to the vortex lines,

$$\oint_{C_1} u\delta x+v\delta y+w\delta z-E\delta t=\oint_{C_2} u\delta x+v\delta y+w\delta z-E\delta t$$

---

* The integral (5) is an integral invariant for other systems as well [for example for the system (9)] for which $dt \neq 0$.

We arbitrarily specify the magnitude $\tau > 0$ and transfer any point of the space (x, y, z, t) to the point $(x', y', z', t + \tau)$ where x', y', z' are the coordinates at the time $t + \tau$ of the fluid particle which at time t had the coordinates x, y, z. In such a displacement along the trajectories of the fluid particles the vortex lines will move to certain new lines which we will call "displaced lines". The tube of vortex lines that we took will move to the tube of displaced lines, and the contours C_1 and C_2 will move into the contours D_1 and D_2 (see Fig. 34).* Since the displacement was accomplished by the motion of fluid particles, the integral (5) does not change its value during the displacement:

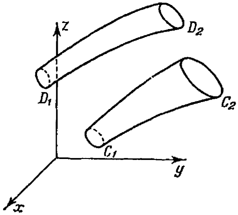

Fig. 34

$$\oint_{C_1} = \oint_{D_1}, \quad \oint_{C_2} = \oint_{D_2}$$

but then

$$\oint_{D_1} = \oint_{D_2} \tag{14}$$

It may be taken that D_1 and D_2 are two arbitrary contours bounding the tube of displaced lines. Therefore, the equation (14) expresses the invariance of the integral (5) with respect to the "displaced lines". Here, $dt = 0$ along each displaced line, as it is along a vortex line. Consequently, displaced lines possess the same properties that (as was shown earlier) only vortex lines can possess. Hence, the displaced lines are vortex lines. And the time of displacement τ is arbitrary. Thus any vortex line always remains a vortex line during the motion of the fluid particles that comprise it.

We have arrived at the *Helmholtz theorem*, which may be formulated as follows: *a vortex line is a fluid line.*

---

* The tubes of vortex lines considered here are located in the four-dimensional space (x, y, z, t) but the t-axis is not depicted in Fig. 34.

At the same time we found that with each vortex tube there is associated a definite "intensity" defined by the integral

$$\oint_C u\delta x + v\delta y + w\delta z - E\delta t \tag{15}$$

The magnitude of this intensity does not change during motion of the fluid. If we take the tube of vortex lines for one and the same instant of time t ($\delta t = 0$), then the intensity (15) is the circulation of the velocity round the contour C:

$$\oint_C u\delta x + v\delta y + w\delta z$$

20. Generalized Conservative Systems. Whittaker's Equations. Jacobi's Equations. The Maupertuis-Lagrange Principle of Least Action

We consider a generalized conservative system, i.e. an arbitrary system (and perhaps a non-natural one as well) for which the function H is not explicitly dependent on the time. For it we have the generalized integral of energy

$$H(q_i, p_i) = h \tag{1}$$

This integral is analogous to the integral of conservation of momentum $p_1 = c$, which holds when q_1 is a cyclic coordinate, i.e. when $\frac{\partial H}{\partial q_1} = 0$.

Proceeding from the analogy between the variable of time t and a cyclic coordinate, we may expect to be able, with the aid of the energy integral (1), to reduce the order of the set of differential equations of motion by two units.

With this purpose in mind, we consider an ordinary (i.e. nonextended) $2n$-dimensional phase space in which the quantities q_i, p_i ($i = 1, \ldots, n$) are the coordinates of a point. We confine ourselves to only those points of the phase space whose coordinates satisfy the equation (1) with fixed value of the constant h_0. In other words, we confine ourselves solely to those states of the system to which the given magnitude of the total energy corresponds:

$$H = H(q_i^0, p_i^0) = h_0$$

Then the basic integral invariant I will appear as

$$I = \oint \sum_{i=1}^{n} p_i \delta q_i, \tag{2}$$

since, by virtue of equality (1),

$$\oint H\delta t = h_0 \oint \delta t = 0$$

Now solve equation (1) for one of the momenta, for example p_1:

$$p_1 = -K(q_1, \ldots, q_n, p_2, \ldots, p_n, h_0) \tag{3}$$

and put the expression obtained into the integral (2) in place of p_1; then

$$I = \oint \Big[\sum_{j=2}^{n} p_j \delta q_j - K\delta q_1\Big] \tag{4}$$

But the integral invariant (4) again has the form of the Poincaré-Cartan integral if it is assumed that the basic coordinates and momenta are the quantities q_j and p_j $(j = 2, \ldots, n)$, and the variable q_1 plays the role of the time variable (in place of the function H we have the function K). For this reason (see Sec. 18), the motion of a generalized conservative system should satisfy the following Hamiltonian system of differential equations of order $2n - 2$:

$$\frac{dq_j}{dq_1} = \frac{\partial K}{\partial p_j}, \quad \frac{dp_j}{dq_1} = -\frac{\partial K}{\partial q_j} \qquad (j = 2, \ldots, n) \tag{5}$$

These equations were obtained by Whittaker and are known as *Whittaker's equations.*

Integrating the system of Whittaker's equations we find the q_j and p_j $(j = 2, \ldots, n)$ as functions of the variable q_1 and $2n - 2$ arbitrary constants $c_1, \ldots, c_{2n-2}$. Moreover, the integrals of Whittaker's equations will contain an arbitrary constant h_0, i.e. they will be of the form

$$\begin{aligned} q_j &= \varphi_j(q_1, h_0, c_1, \ldots, c_{2n-2}), \\ p_j &= \psi_j(q_1, h_0, c_1, \ldots, c_{2n-2}) \qquad (j = 2, \ldots, n). \end{aligned} \tag{6}$$

Then by putting expression (6) into formula (3) we find an analogous expression for

$$p_1 = \psi_1(q_1, h_0, c_1, \ldots, c_{2n-2}) \tag{6'}$$

All the trajectories in phase space (i.e. the geometrical pattern of motion) are determined in this way.

We obtain the dependence of the coordinates on the time t from the equation $\frac{dq_1}{dt} = \frac{\partial H}{\partial p_1}$ by means of a quadrature:

$$t = \int \frac{dq_1}{\frac{\partial H}{\partial p_1}} + c_{2n-1} \tag{7}$$

where all the variables in the partial derivative $\frac{\partial H}{\partial p_1}$ are expressed in terms of q_1 with the aid of the equations (6) and (6′).

The Hamiltonian system of Whittaker's equations (5) may be replaced by an equivalent system of equations of the Lagrangian type:

$$\frac{d}{dq_1}\frac{\partial P}{\partial q_j'}-\frac{\partial P}{\partial q_j}=0 \qquad (j=2, \ldots, n) \tag{8}$$

where $q_j'=\frac{dq_j}{dq_1}$, and the function P (the analogue of the Lagrangian function) is connected with the function K (the analogue of the Hamiltonian function) by the equation*

$$P=P(q_1, \ldots, q_n, q_2', \ldots, q_n')=\sum_{j=2}^{n} p_j q_j'-K \tag{9}$$

Note that the system (8) contains not n but $n-1$ second-order equations. In the latter part of this relation the momenta p_j must be replaced by their expression in terms of

$$q_2'=\frac{dq_2}{dq_1}, \ \ldots, \ q_n'=\frac{dq_n}{dq_1},$$

which may be obtained from the first $n-1$ equations (5).

We transform the expression for the function P by proceeding from the equations (3) and (9):

$$P=\sum_{j=2}^{n} p_j q_j'+p_1=\frac{1}{\dot q_1}\sum_{i=1}^{n} p_i \dot q_i=\frac{1}{\dot q_1}(L+H) \tag{10}$$

In the case of a natural, i.e. an ordinary conservative system, $L=T-\Pi$ and $H=T+\Pi$; consequently,

$$P=\frac{1}{\dot q_1}(L+H)=\frac{2T}{\dot q_1} \tag{11}$$

and the kinetic energy T may be written as

$$T=\frac{1}{2}\sum_{i,k=1}^{n} a_{ik}\dot q_i \dot q_k=\dot q_1^2 G(q_1, \ldots, q_n, q_2', \ldots, q_n') \tag{12}$$

where

$$G(q_1, \ldots, q_n, q_2', \ldots, q_n')=\frac{1}{2}\sum_{i,k=1}^{n} a_{ik} q_i' q_k' \quad (q_1'=1) \tag{13}$$

---

* See equation (8) on page 73.

From equations (1) and (12) we find

$$\dot{q}_1 = \sqrt{\frac{h-\Pi}{G}} \tag{14}$$

and for the function P we get the following expression:

$$P = \frac{2T}{\dot{q}_1} = 2\sqrt{G(h-\Pi)} \tag{15}$$

The differential equations (8), in which the function P is of the form (15) and which therefore belong to the natural (i.e. ordinary) conservative systems are referred to as Jacobi equations.*

Integrating Jacobi's equations, we determine all the trajectories in the coordinate space $(q_1, \ldots, q_n)$:

$$q_j = \varphi_j(q_1, h, c_1, \ldots, c_{2n-2}) \tag{16}$$

The relation between the coordinates and the time variable is established from (14) by a quadrature:

$$t = \int \sqrt{\frac{G}{h-\Pi}}\, dq_1 + c_{2n-1} \tag{17}$$

Let us now establish the principle of least action of Maupertuis-Lagrange. Since the equations (8) are Lagrangian type equations, there follows the variational principle according to which for the straight-line path

$$\delta W^* = 0 \tag{18}$$

where

$$W^* = \int_{q_1^0}^{q_1^1} P\, dq_1 \tag{19}$$

is the Lagrange action. Here, all the motions of a generalized-conservative system that transfer the system from a given initial position q_1^0 to a specified terminal position q_1^1 (Fig. 35) are permitted to contend. The instants of time t_0 and t_1 are not fixed and may vary when passing from the straight-line path to circuitous paths.**

Expressing the function P in the integral (19) with the aid of equation (10), we find the connection between Lagrange action W^*

---

* These equations were derived by the German mathematician K.G.J. Jacobi and are found in his classical *Vorlesungen über Dynamik* published in 1886. In the case of a non-natural generalized-conservative system, the function P involved in the Jacobi equations is determined by the formula (9).

** Equation (18) states that for a straight-line path, the Lagrange action has a fixed value. The question of when the action W^* is least for a straight-line path is solved by invoking kinetic foci in the same way as for Hamilton's principle.

and Hamilton action W:

$$W^* = \int_{t_0}^{t_1} (L+H)\,dt = \int_{t_0}^{t_1} L\,dt + h\int_{t_0}^{t_1} dt = W + h\,(t_1 - t_0) \qquad (20)$$

For an ordinary (natural) conservative system, it is possible to take advantage of expression (11) and represent the Lagrange action in the form

$$W^* = 2\int_{t_0}^{t_1} T\,dt = \int_{t_0}^{t_1} \sum_{\nu=1}^{N} m_\nu v_\nu^2\,dt = \sum_{\nu=1}^{N} \int_{s_\nu^0}^{s_\nu^1} m_\nu v_\nu\,ds_\nu \qquad (21)$$

But it has already been assumed that only those motions of a system are permitted to "compete", for which the total energy of the system has one and the same value of h.

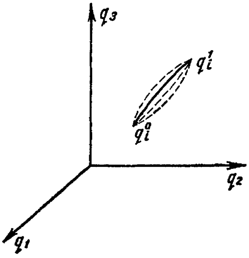

Fig. 35

The expression obtained for W^* indicates that the Lagrange action W^* is equal to the work of the momentum vectors of points of the system in a corresponding displacement of the system.

The variational principle (18) is referred to as the *Maupertuis-Lagrange principle of least action.**

By way of illustration let us compare the optical principle of Fermat and the principle of least action of Maupertuis-Lagrange **

---

* First formulated vaguely by Maupertuis in 1747. In 1760 Lagrange gave a rigorous formulation and substantiation of the principle.

** See, for example, H. V. Rose *Lectures in Analytical Mechanics*, Part 1. Leningrad, 1938, pp. 93-94 (in Russian).

According to Fermat's principle, light is propagated in an inhomogeneous medium so that the time of flight

$$t=\int_{s_0}^{s}\frac{ds}{v} \tag{22}$$

is a minimum.

Introducing at each point of the medium the refractive index $n=\frac{c}{v}$, i.e. the ratio of the speed of light c in vacuo to the speed v in the given medium, we transform the formula to

$$t=\frac{1}{c}\int_{s_0}^{s} n\,ds$$

where n is a function of a point of the medium. Then the Fermat principle is written as

$$\delta\int_{s_0}^{s} n\,ds=0 \tag{23}$$

On the other hand, for a single particle of mass m the Maupertuis-Lagrange principle yields $\left(\text{since } \frac{1}{2}mv^2+\Pi=h\right)$

$$\delta\int_{s_0}^{s} mv\,ds=\sqrt{m}\,\delta\int_{s_0}^{s}\sqrt{2(h-\Pi)}\,ds=0 \tag{24}$$

The trajectory of a light ray coincides with the trajectory of a particle if [see formulas (23) and (24)]

$$\Pi=h-\frac{1}{2}n^2 \tag{25}$$

If it is assumed that near the earth's surface the refractive index n diminishes as a linear function of the altitude z

$$n=n_0\left(1-k\frac{z}{H}\right) \tag{26}$$

where H is the height of the atmosphere, then, neglecting small terms of the order of $\left(\frac{z}{H}\right)^2$, we can write

$$\Pi=h-\frac{1}{2}n_0^2\left(1-2k\frac{z}{H}\right)=c+gz \tag{27}$$

where

$$c=h-\frac{1}{2}n_0^2,\qquad g=\frac{kn_0^2}{H} \tag{28}$$

Formula (27) defines the potential of the force of gravity near the surface of the earth with a modified value of g.

Therefore, if the refractive index n varies with altitude in the manner indicated, light will be propagated in a parabola with a vertical axis.

21. Inertial Motion. Relation to Geodesic Lines in the Arbitrary Motion of a Conservative System

Let there be given an arbitrary scleronomic system; its kinetic energy is

$$T=\frac{1}{2}\sum_{i,k=1}^{n} a_{ik}(q_1, \ldots, q_n)\,\dot{q}_i\dot{q}_k \tag{1}$$

We introduce a metric in the coordinate space $(q_1, \ldots, q_n)$ by determining the square of the arc length ds^2 with the aid of the positive-definite quadratic differential form

$$ds^2=\sum_{i,k=1}^{n} a_{ik}(q_1, \ldots, q_n)\,dq_i\,dq_k \tag{2}$$

Then the magnitude of the arc of the curve connecting two points of coordinate space, (q_i^0) and (q_i), will be determined by the equation

$$s=\int_{(q_i^0)}^{(q_i)} ds=\int_{(q_i^0)}^{(q_i)} \sqrt{\sum_{i,k=1}^{n} a_{ik}\,dq_i\,dq_k} \tag{3}$$

Comparing formulas (1) and (2) we find that in the motion of the system

$$T=\frac{1}{2}\left(\frac{ds}{dt}\right)^2 \tag{4}$$

That is, *the kinetic energy of a system* [given the metric (2)] *always coincides with the kinetic energy of the representative point in the n-dimensional coordinate space if to this point we assign a mass of* $m = 1$.

Now consider the inertial motion of the system ($\Pi = 0$). Then in that motion all possible trajectories of the image point are referred to as *geodesic lines* [with respect to the metric (2)]. From the energy integral

$$T=h$$

by formula (4), it follows that

$$\frac{ds}{dt}=\sqrt{2h} \tag{5}$$

That is, to inertial motion (and also to any motion with a constant value h of the kinetic energy) there corresponds in the coordinate space $(q_1, \ldots, q_n)$ a uniform motion of the representative point with velocity $\sqrt{2h}$.

In accordance with the principle of least action, the geodesic lines are extremals* of the variational problem

$$\delta W^* = 0 \tag{6}$$

where W^* is the Lagrange action. However, in the case under consideration we have the energy integral $T = h$ with fixed value of h both for the straight-line path and for a circuitous path. Therefore,

$$W^* = \int_{t_0}^{t} 2T\,dt = 2h(t - t_0) = \sqrt{2h}\,s \tag{7}$$

where $s = \sqrt{2h}(t - t_0)$ is the length of the curve in the coordinate space $(q_1, \ldots, q_n)$.

The variational problem (6) takes on the form

$$\delta s = \delta \int_{q_i^0}^{q_i} \sqrt{\sum_{i,k=1}^{n} a_{ik}\,dq_i\,dq_k} = 0 \tag{8}$$

Thus a geodesic line is distinguished by the fact that the arc length of the curve has an extremal (stationary, to be more precise) value as compared with the arcs of other curves having the same end points as the geodesic line** (see Fig. 35 on page 115).

When for a straight-line path the Lagrange action is a minimum, the arc length of the geodesic is less than the length of any other curve connecting the same points as does the arc of the geodesic. That is why geodesic lines are also called the *shortest lines* in space.

Now consider a conservative system, i.e. a scleronomic system with potential energy $\Pi = \Pi(q_1, \ldots, q_n)$ not explicitly dependent on t. Then, by (15) and (19) of the preceding section,

$$W^* = 2\int_{q_1^0}^{q_1} \sqrt{G(h - \Pi)}\,dq_1 = 2\int_{q_i^0}^{q_i} \sqrt{(h - \Pi)\sum_{i,k=1}^{n} a_{ik}\,dq_i\,dq_k} \tag{9}$$

Therefore, the motion of a conservative system with a given value of total energy h is accomplished in a coordinate space along the extremal of the variational problem (with fixed end points):

$$\delta \int_{(q_i^0)}^{(q_i)} \sqrt{(h - \Pi)\sum_{i,k=1}^{n} a_{ik}\,dq_i\,dq_k} = 0 \tag{10}$$

---

* That is, curves for which W^* has a stationary value.

** This always occurs when the ends of the arc of a geodesic line are sufficiently close together.

Comparing (10) and (8) we conclude that for a conservative system the trajectories of straight-line paths are geodesic lines in a coordinate space with the metric

$$ds_1^2 = (h - \Pi) \sum_{i,k=1}^{n} a_{ik}\, dq_i\, dq_k \qquad (11)$$

22. The Universal Integral Invariant of Poincaré. Lee Hwa-Chung's Theorem

Now consider the integral $I = \oint \left[\sum_{i=1}^{n} p_i \delta q_i - H\delta t \right]$ round the contour C consisting of simultaneous states of a system. Such a contour results if a tube of straight-line paths (see Fig. 33 on page 100) is cut by a hyperplane $t = \text{const}$. For such a contour, $\delta t = 0$ and the basic integral invariant takes the form

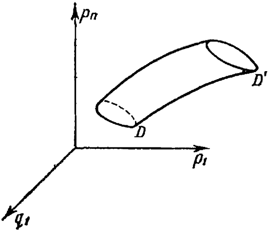

Fig. 36

$$I_1 = \oint \sum_{i=1}^{n} p_i \delta q_i \qquad (1)$$

This integral was first introduced by Poincaré. Later, Cartan extended the integral to contours consisting of nonsimultaneous states by introducing an additional summand $-H\delta t$.

Poincaré's integral invariant I_1 does not change its value if the contour C is displaced along the tube of straight-line paths to the contour C' which again consists of simultaneous states. It is convenient to consider the integral I_1 in the ordinary (nonextended) $2n$-dimensional phase space $(q_1, p_1, \ldots, q_n, p_n)$. In this space, to the contours C and C' (Fig. 33) there correspond the contours D and D'

bounding the tube of "straight-line" paths (Fig. 36). Here,

$$\oint_D \sum_{i=1}^{n} p_i \delta q_i = \oint_{D'} \sum_{i=1}^{n} p_i \delta q_i$$

Note that one of the contours D and D' (say D) may be chosen quite arbitrarily. It may be taken that the points of the contour D are different states of the system at one and the same instant of time t; then the corresponding states of the system at time t' will form the contour D'.

In place of the phase space we consider n separate phase planes (q_i, p_i) $(i = 1, \ldots, n)$. Project an arbitrary closed curve (contour) D situated in the phase space onto these planes (Fig. 37). We get the contours D_i $(i = 1, \ldots, n)$. Then for any i

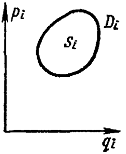

Fig. 37

$$\oint_D p_i \delta q_i = \oint_{D_i} p_i \delta q_i = \pm S_i \tag{2}$$

where S_i is the area of the region bounded by the contour D_i in the plane (q_i, p_i) $(i = 1, \ldots, n)$. The sense of circulation about the contour D induces the sense of circulation on the projection D_i. In formula (2) the sign before S_i is plus if the circulation of D_i is clockwise (that is, in the direction of the shortest turn of the axis p_i to the axis q_i), and it is minus otherwise.

Then

$$I_1 = \oint \sum_{i=1}^{n} p_i \delta q_i = \sum_{i=1}^{n} \pm S_i \tag{3}$$

Thus the contours D and D_i vary during the motion of the system, and the areas S_i also vary, but the algebraic sum (3) of these areas remains constant. That is the geometrical interpretation of the invariance of the Poincaré integral.

H does not appear in the expression for I_1. Consequently, Poincaré's integral I_1 is invariant with respect to any Hamiltonian system. That is why the integral I_1 is called the *universal integral invariant.*

The following proposition may likewise be readily proved.

If for some system of differential equations

$$\frac{dq_i}{dt}=Q_i(t, q_k, p_k), \quad \frac{dp_i}{dt}=P_i(t, q_k, p_k) \qquad (i=1, \ldots, n) \tag{4}$$

the integral I_1 is invariant, then the system (4) is Hamiltonian.

Indeed, in this case

$$0=\frac{d}{dt}I_1=\oint\sum_{i=1}^{n}\left(\frac{dp_i}{dt}\delta q_i+p_i\frac{d}{dt}\delta q_i\right)=$$

$$=\oint\sum_{i=1}^{n}\left(\frac{dp_i}{dt}\delta q_i+p_i\delta\frac{dq_i}{dt}\right)=\oint\sum_{i=1}^{n}\left(\frac{dp_i}{dt}\delta q_i-\frac{dq_i}{dt}\delta p_i\right)=$$

$$=\oint\sum_{i=1}^{n}(P_i\delta q_i-Q_i\delta p_i)$$

From this it follows that the expression under the integral sign is a total virtual differential of some function $-H(t, q_k, p_k)$:*

$$\sum_{i=1}^{n}(P_i\delta q_i-Q_i\delta p_i)=-\delta H(t, q_k, p_k)$$

that is

$$Q_i=\frac{\partial H}{\partial p_i}, \quad P_i=-\frac{\partial H}{\partial q_i} \qquad (i=1, \ldots, n)$$

and the proof is complete.

Note the following terms: the Poincaré-Cartan integral I and the Poincaré integral I_1 are called *relative* integral invariants of the first order. The term 'relative' means that the domain of integration is a closed contour and the 'first order' implies that the differentials enter linearly in the expression under the integral sign. Note that the relative integral invariant of the first order I_1 may, by means of Stokes' formula, be represented in the form of an absolute integral invariant of the second order:

$$\oint_D\sum_{i=1}^{n}A_i\delta x_i=\iint_S\sum_{i<k}^{1,\ldots,n}\left(\frac{\partial A_k}{\partial x_i}-\frac{\partial A_i}{\partial x_k}\right)\delta x_i\delta x_k \tag{5}$$

---

* Here, $\delta H=\sum_{i=1}^{n}\left(\frac{\partial H}{\partial q_i}\delta q_i+\frac{\partial H}{\partial p_i}\delta p_i\right)$.

where the integral on the right is taken over the surface S bounded by the closed contour D.

This formula, as applied to the integral I_1, yields

$$I_1 = \iint \sum_{i=1}^{n} \delta p_i \delta q_i$$

In $2n$-dimensional phase space there are known to exist the following universal relative integral invariants I_{2k-1} of odd orders and the absolute integral invariants J_{2k} of even orders,

$$I_1 = \oint \sum p_i \delta q_i = J_2 = \iint \sum \delta p_i \delta q_i$$

$$I_3 = \oint\!\!\oint\!\!\oint \sum p_i \delta q_i \delta p_k \delta q_k = J_4 = \int\!\!\int\!\!\int\!\!\int \sum \delta p_i \delta q_i \delta p_k \delta q_k$$

. . . .

$$I_{2n-1} = \oint \cdots \oint \sum p_{i_1} \delta q_{i_1} \ldots \delta p_{i_n} \delta q_{i_n} =$$

$$= J_{2n} = \int \ldots \int \delta p_{i_1} \delta q_{i_1} \ldots \delta p_{i_n} \delta q_{i_n}$$

In 1947, the Chinese scientist Lee Hwa-Chung proved the uniqueness of these universal integral invariants. He demonstrated that any other universal integral invariant differs by a constant factor from one of the enumerated integrals.

We shall need the Lee Hwa-Chung theorem later on for an integral invariant of the first order; we therefore formulate the theorem for an arbitrary n and will prove it for $n = 1$.

Theorem of Lee Hwa-Chung. *If*

$$I' = \oint \sum_{i=1}^{n} [A_i(t, q_k, p_k) \delta q_i + B_i(t, q_k, p_k) \delta p_i]$$

is a universal relative integral invariant, then

$$I' = cI_1 \tag{6}$$

where c is a constant, and I_1 is the Poincaré integral.

Proof. Let

$$I' = \oint A(t, q, p) \delta q + B(t, q, p) \delta p$$

be a universal integral invariant. The integration is taken round a closed contour in the phase plane (q, p). Further, let there be given some Hamiltonian system of differential equations with the function $H(t, q, p)$:

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}, \quad \frac{dp}{dt} = -\frac{\partial H}{\partial q} \tag{7}$$

The general solution of this system is of the form

$$q = q(t, q_0, p_0), \quad p = p(t, q_0, p_0) \tag{8}$$

where q_0, p_0 are the initial values of q and p for $t = t_0$.

Let

$$q = q_0(\alpha), \quad p = p_0(\alpha) \tag{9}$$

$$[0 \leqslant \alpha \leqslant l;\ q_0(0) = q_0(l),\ p_0(0) = p_0(l)]$$

be the equations of the closed contour D_0 in the phase plane (Fig. 38). The points which at the time $t = t_0$ were located on the contour D_0 form the contour D at some other arbitrary instant of time t. The parametric equations of this contour are obtained from the equalities (8) if q_0 and p_0 are replaced by their expressions (9). This yields

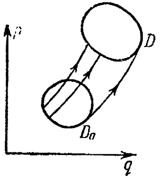

Fig. 38

$$q = q(t, \alpha), \quad p = p(t, \alpha) \quad (0 \leqslant \alpha \leqslant l) \tag{10}$$

Putting these functions in the integral I' in place of q and p, we get I' as a function of the parameter t. From the invariance of I' it follows that $\frac{dI'}{dt} = 0$. Differentiating under the integral sign and integrating by parts* we find

$$0 = \frac{dI'}{dt} = \oint \frac{dA}{dt}\,\delta q + \frac{dB}{dt}\,\delta p + A\frac{d}{dt}\,\delta q + B\frac{d}{dt}\,\delta p =$$

$$= \oint \frac{dA}{dt}\,\delta q + \frac{dB}{dt}\,\delta p + A\delta\frac{dq}{dt} + B\delta\frac{dp}{dt} =$$

$$= \oint \frac{dA}{dt}\,\delta q - \delta A\frac{dq}{dt} + \frac{dB}{dt}\,\delta p - \delta B\frac{dp}{dt} =$$

$$= \oint \left[\left(\frac{\partial A}{\partial p} - \frac{\partial B}{\partial q}\right)\frac{dp}{dt} + \frac{\partial A}{\partial t}\right]\delta q + \left[\left(\frac{\partial B}{\partial q} - \frac{\partial A}{\partial p}\right)\frac{dq}{dt} + \frac{\partial B}{\partial t}\right]\delta p =$$

$$= \oint \left(-Z\frac{\partial H}{\partial q} + \frac{\partial A}{\partial t}\right)\delta q + \left(-Z\frac{\partial H}{\partial p} + \frac{\partial B}{\partial t}\right)\delta p$$

---

* See the footnote * on page 103. When making transformations, we assume $\delta A = \frac{\partial A}{\partial q}\,\delta q + \frac{\partial A}{\partial p}\,\delta p$ and $\delta B = \frac{\partial B}{\partial q}\,\delta q + \frac{\partial B}{\partial p}\,\delta p$ and then use the equations (7).

where

$$Z = \frac{\partial A}{\partial p} - \frac{\partial B}{\partial q}$$

The last integral is equal to zero for any value of the variable t considered as a parameter and for an arbitrary path of integration. Therefore the expression under the sign of the integral must be a total differential with respect to the variables q and p. From this,

$$\frac{\partial}{\partial p}\left(-Z\frac{\partial H}{\partial q} + \frac{\partial A}{\partial t}\right) = \frac{\partial}{\partial q}\left(-Z\frac{\partial H}{\partial p} + \frac{\partial B}{\partial t}\right)$$

or, after elementary transformations,

$$-\frac{\partial Z}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial Z}{\partial q}\frac{\partial H}{\partial p} + \frac{\partial Z}{\partial t} = 0$$

Since the function H may be chosen quite arbitrarily, it follows that

$$\frac{\partial Z}{\partial p} = \frac{\partial Z}{\partial q} = \frac{\partial Z}{\partial t} = 0$$

that is,

$$Z = \frac{\partial A}{\partial p} - \frac{\partial B}{\partial q} = \text{const} = c$$

Then

$$\frac{\partial (A - cp)}{\partial p} = \frac{\partial B}{\partial q}$$

and, consequently, there exists a function $\Phi(t, q, p)$ such that*

$$(A - cp)\,\delta q + B\delta p = \frac{\partial \Phi}{\partial q}\delta q + \frac{\partial \Phi}{\partial p}\delta p = \delta\Phi$$

But then

$$A\delta q + B\delta p = cp\delta q + \delta\Phi$$

and therefore

$$I' = \oint A\delta q + B\delta p = c\oint p\delta q = cI_1$$

which completes the proof.

For $n > 1$, the idea of proof is retained, though the proof itself becomes more involved.

---

* Here, the time t is taken as a parameter.

23. Invariance of Volume in the Phase Space. Liouville's Theorem

We consider the "total" absolute integral invariant

$$J = \int \dots \int \delta p_1 \delta q_1 \dots \delta p_n \delta q_n \tag{1}$$

The invariance of this integral implies the invariance of the phase volume in a $2n$-dimensional phase space and is established in the following manner.*

Let us write down in the following form the final equations of motion that result after integration of Hamilton's equations:

$$q_i = q_i(t, q_k^0, p_k^0), \quad p_i = p_i(t, q_k^0, p_k^0)$$
$$(i = 1, \dots, n) \tag{2}$$

where q_k^0, p_k^0 are the initial values of q_k, p_k for $t = t_0$ $(k = 1, \dots, n)$.

In phase space we choose some volume J_0 and take each point of this volume as the initial point (for $t = t_0$). Then by time t the transformation (2) will have carried the volume J_0 into the volume J. Here,

$$J = \underbrace{\int \dots \int}_{J_0} I \delta q_1^0 \delta p_1^0 \dots \delta q_n^0 \delta p_n^0 \tag{3}$$

where

$$I = \left| \frac{\partial(q_1, p_1, \dots, q_n, p_n)}{\partial(q_1^0, p_1^0, \dots, q_n^0, p_n^0)} \right|$$

that is, I is the Jacobian composed of the partial derivatives of q_j, p_j with respect to the initially given q_i^0, p_i^0 $(i, j = 1, \dots, n)$. Without loss of generality, we can consider that the Jacobian under the integral sign in (3) is positive and we can drop the absolute-value sign.

When $t = t_0$, this Jacobian is equal to 1 since for this value of t all the $q_k = q_k^0$ and the $p_k = p_k^0$. When t varies, the Jacobian varies continuously without vanishing, since the singular points at which this Jacobian might vanish are eliminated from our consideration; i. e. it is assumed that there are no such points in the volume under examination. Then the Jacobian is positive in this volume.**

---

* Henceforward (see Sec. 31) the invariance of the phase volume will be established on the basis of the general properties of motion of Hamiltonian systems.

** No assumptions concerning singular points are required in the proof on pages 161-162 mentioned in the preceding footnote.

Differentiating with respect to t under the integral sign we get

$$\left(\frac{dJ}{dt}\right)_{t=t_0}=\underbrace{\int\cdots\int}_{J_0}\left|\frac{dI}{dt}\right|_{t=t_0}\delta q_1\delta p_1\ldots\delta q_n\delta p_n$$

Compute the derivative of the Jacobian I:

$$\left|\frac{dI}{dt}\right|_{t=t_0}=\sum_{i=1}^{2n}|I_i|_{t=t_0}$$

where I_i is the determinant obtained from the Jacobian by differentiation of the ith row. Now taking into account that

$$1^\circ \text{ for } i\neq j\ \left|\frac{\partial q_j}{\partial q_i^0}\right|_{t=t_0}=\left|\frac{\partial p_j}{\partial p_i^0}\right|_{t=t_0}=0$$

$$2^\circ \text{ for any } i\ \left|\frac{\partial q_i}{\partial p_i^0}\right|_{t=t_0}=\left|\frac{\partial p_i}{\partial q_i^0}\right|_{t=t_0}=0$$

$$3^\circ \text{ for any } i\ \left|\frac{\partial q_i}{\partial q_i^0}\right|_{t=t_0}=\left|\frac{\partial p_i}{\partial p_i^0}\right|_{t=t_0}=1$$

we find

$$|I_i|_{t=t_0}=\left|\frac{\partial \dot{q}_i}{\partial q_i^0}\right|_{t=t_0}\quad \text{for } i=1,\ \ldots,\ n$$

and

$$|I_i|_{t=t_0}=\left|\frac{\partial \dot{p}_i}{\partial p_i^0}\right|_{t=t_0}\quad \text{for } i=n+1,\ \ldots,\ 2n$$

therefore

$$\left|\frac{dI}{dt}\right|_{t=t_0}=\sum_{i=1}^{2n}|I_i|_{t=t_0}=\sum_{i=1}^{n}\left[\frac{\partial \dot{q}_i}{\partial q_i^0}+\frac{\partial \dot{p}_i}{\partial p_i^0}\right]_{t=t_0}=$$

$$=\sum_{i=1}^{n}\left[\frac{\partial}{\partial q_i^0}\frac{\partial H_0}{\partial p_i^0}-\frac{\partial}{\partial p_i^0}\frac{\partial H_0}{\partial q_i^0}\right]=\sum_{i=1}^{n}\left[\frac{\partial^2 H_0}{\partial q_i^0\,\partial p_i^0}-\frac{\partial^2 H_0}{\partial p_i^0\,\partial q_i^0}\right]=0$$

Thus $\left(\frac{dJ}{dt}\right)_{t=t_0}=0$. Since the initial instant t_0 may be chosen quite arbitrarily, it follows that for any t

$$\frac{dJ}{dt}=0 \tag{4}$$

which is to say that the magnitude of the phase volume J does not change upon a displacement of the points of this volume from states occupied at time t_0 to states occupied at any other arbitrary instant of time t.

From the invariance of the phase volume there follows one of the basic theorems of statistical mechanics, Liouville's theorem.

Imagine a very large number of absolutely identical replicas of a system differing from one another solely in their initial states q_i^0, p_i^0 $(i = 1, \ldots, n)$. All these replicas form a *statistical ensemble*.

An instance of a statistical ensemble is the collection of molecules of a gas in a given volume.

To each volume element dV of the phase space we can assign a "mass" $d\mu$ that describes the quantity of replicas per given element of volume dV. By virtue of the proved invariance of volume in phase space, the magnitude of dV does not change with time. Physically speaking, the magnitude of $d\mu$ does not change either, since the replicas which at some time are in the volume dV will be displaced together with the volume. That is why the density of the statistical ensemble remains unchanged during motion:

$$\rho(t, q_i, p_i) = \frac{d\mu}{dV} \tag{5}$$

i.e.

$$\frac{d\rho}{dt} = 0 \tag{6}$$

In expanded form, equation (5) may be written thus (see Sec. 15):

$$\frac{\partial\rho}{\partial t} + (\rho H) = 0 \tag{6'}$$

where (ρH) is a Poisson bracket. According to (6), the function $\rho(t, q_i, p_i)$ is the integral of motion.

We have thus proved the following theorem.

Liouville's theorem. *The density of a statistical ensemble is always an integral of motion.*

Thus, for example, for a conservative system any function of the energy of the system may serve as the density of the statistical ensemble.

CHAPTER 4

Canonical Transformations and the Hamilton-Jacobi Equation

24. Canonical Transformations

The transformation of coordinates in a $2n$-dimensional phase space (containing, in the general case, the time variable t as a parameter)

$$\begin{aligned}\tilde{q}_i &= \tilde{q}_i(t, q_k, p_k) \\ \tilde{p}_i &= \tilde{p}_i(t, q_k, p_k)\end{aligned} \quad \left(i = 1, \ldots, n;\ \frac{\partial(\tilde{q}_1, \tilde{p}_1, \ldots, \tilde{q}_n, \tilde{p}_n)}{\partial(q_1, p_1, \ldots, q_n, p_n)} \neq 0\right) \tag{1}$$

is called *canonical* if the transformation carries any Hamiltonian system

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad (i = 1, \ldots, n) \tag{2}$$

again into a Hamiltonian system (generally speaking, with another Hamiltonian function $\tilde{H}$):

$$\frac{d\tilde{q}_i}{dt} = \frac{\partial \tilde{H}}{\partial \tilde{p}_i}, \quad \frac{d\tilde{p}_i}{dt} = -\frac{\partial \tilde{H}}{\partial \tilde{q}_i} \quad (i = 1, \ldots, n) \tag{3}$$

The importance of studying canonical transformations is due to the fact that these transformations permit replacing a given Hamiltonian system (2) by another Hamiltonian system (3) in which the function $\tilde{H}$ is of a simpler structure than H.

If in a phase space we perform two canonical transformations in succession, the resulting transformation will again be canonical. Furthermore, a transformation that is inverse to a certain canonical transformation will always be canonical, and the identical transformation $\tilde{q}_i = q_i$, $\tilde{p}_i = p_i$ $(i = 1, \ldots, n)$ is canonical. Therefore, all canonical transformations taken together form a *group*.

Example 1. The transformation

$$\tilde{q}_i = \alpha q_i, \ \tilde{p}_i = \beta p_i \quad (i = 1, \ldots, n;\ \alpha \neq 0,\ \beta \neq 0)$$

as may readily be verified is canonical. It transforms the system (2) into the system (3) with

$$\tilde{H} = \alpha\beta H$$

Example 2. The transformation

$$\tilde{q}_i = \alpha p_i, \ \tilde{p}_i = \beta q_i \quad (i = 1, \ldots, n;\ \alpha \neq 0,\ \beta \neq 0)$$

will be canonical. In this case,

$$\tilde{H} = -\alpha\beta H$$

Example 3. The transformation

$$\tilde{q}_i = p_i \tan t, \quad \tilde{p}_i = q_i \cot t \quad (i = 1, \ldots, n)$$

will be canonical because it is readily verifiable that from the equations (2) we always get equations (3) for

$$\tilde{H} = -H + \frac{1}{\sin t \cos t} \sum_{i=1}^{n} \tilde{q}_i \tilde{p}_i$$

To derive the conditions under which the transformation (1) is canonical, we consider two extended $(2n+1)$-dimensional phase spaces (q_i, p_i, t) and $(\tilde{q}_i, \tilde{p}_i, t)$, one passing into the other under the canonical transformation (1), and two tubes of straight-line paths of Hamiltonian systems (2) and (3) (Fig. 39).

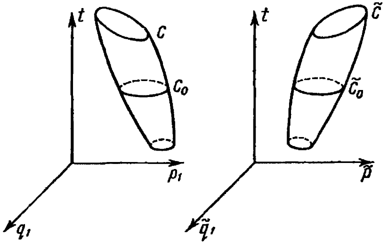

Fig. 39

Let us take two arbitrary closed contours C and $\tilde{C}$ that bound these tubes and correspond to one another by virtue of the transformation (1). Furthermore, cut both tubes with one and the same hyperplane $t = \text{const}$. In the cross-section we get two "plane" contours C_0 and $\tilde{C}_0$. These contours likewise pass one into the other under the canonical transformation (1), since the quantity t remains unchanged in a canonical transformation. From the invariance of the Poincaré-Cartan integral it follows that

$$\oint_C \Big[\sum_{i=1}^{n} p_i \delta q_i - H\delta t\Big] = \oint_{C_0} \sum_{i=1}^{n} p_i \delta q_i \tag{4}$$

$$\oint_{\tilde{C}} \Big[\sum_{i=1}^{n} \tilde{p}_i \delta \tilde{q}_i - \tilde{H}\delta t\Big] = \oint_{\tilde{C}_0} \sum_{i=1}^{n} \tilde{p}_i \delta \tilde{q}_i \tag{5}$$

On the other hand, if in the universal integral invariant $\oint \sum_{i=1}^{n} \tilde{p}_i \delta \tilde{q}_i$ we pass to the variables q_i, p_i $(i = 1, \ldots, n)$ by means of the canonical transformation (1), then this integral will pass into a certain universal integral invariant of the first order in $2n$-dimensional phase (q_i, p_i)-space; by the theorem of Lee Hwa-Chung, the invariant obtained can differ from $\oint \sum_{i=1}^{n} p_i \delta q_i$ only in the constant factor c. Therefore

$$\oint_{\tilde{C}_0} \sum_{i=1}^{n} \tilde{p}_i \delta \tilde{q}_i = c \oint_{C_0} \sum_{i=1}^{n} p_i \delta q_i \tag{6}$$

From the equations (4) to (6) it follows that

$$\oint_{\tilde{C}} \left[\sum_{i=1}^{n} \tilde{p}_i \delta \tilde{q}_i - \tilde{H} \delta t \right] = c \oint_{C} \left[\sum_{i=1}^{n} p_i \delta q_i - H \delta t \right] \tag{7}$$

If in the first integral we consider the variables $\tilde{q}_1, \ldots, \tilde{p}_n$ to be expressed in terms of the variables $q_1, \ldots, p_n$ (here, the path of integration $\tilde{C}$ is replaced by the path of integration C), then the equation (7) may be written as

$$\oint_{C} \left[\sum_{i=1}^{n} \tilde{p}_i \delta \tilde{q}_i - \tilde{H} \delta t \right] - c \left[\sum_{i=1}^{n} p_i \delta q_i - H \delta t \right] = 0 \tag{8}$$

But C is an absolutely arbitrary contour in a $(2n+1)$-dimensional extended phase space. Therefore the expression under the integral sign in (8) must be a total differential of some function of the $2n+1$ arguments $q_1, p_1, \ldots, q_n, p_n$ and t. We will find it convenient to denote this function in terms of $-F\,(t, q_i, p_i)$; then*

$$\sum_{i=1}^{n} \tilde{p}_i \delta \tilde{q}_i - \tilde{H} \delta t = c \left(\sum_{i=1}^{n} p_i \delta q_i - H \delta t \right) - \delta F \tag{9}$$

Note that the constant c in the identity (9) is always different from zero, $c \neq 0$, since the expression $\sum_{i=1}^{n} \tilde{p}_i \delta \tilde{q}_i - \tilde{H} \delta t$ is not a total differential** and therefore cannot be equal to $-\delta F$.

---

* Here, $\delta F = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q_i} \delta q_i + \frac{\partial F}{\partial p_i} \delta p_i \right) + \frac{\partial F}{\partial t} \delta t$.

** With respect to the independent variables $\tilde{q}_i$, $\tilde{p}_i$, t and, hence, with respect to the independent variables q_i, p_i, t $(i = 1, \ldots, n)$.

We shall call the function F the *generating function*, and the constant c the *valence* of the canonical transformation (1) under consideration. The canonical transformation will be called *univalent* if for it $c = 1$.

A necessary and sufficient condition for the transformation (1) *to be canonical is the existence of a generating function F and some constant c for which the equation* (9) *is identically satisfied by virtue of the transformation* (1).

Note. If the transformation (1) is canonical, then there exist a generating function F and a valence $c \neq 0$ such that the equation (9) holds for *any* function H and the corresponding function $\tilde{H}$. However, if (9) holds for one pair of functions H and $\tilde{H}$, then transformation (1) is already canonical.*

Indeed, in addition to H take another arbitrary function H_1 and define $\tilde{H}_1$ from the condition

$$\tilde{H}_1 - \tilde{H} = c\,(H_1 - H)$$

Multiplying both parts of this equation by δt and subtracting termwise the resulting equation from (9), we get

$$\sum_{i=1}^{n} \tilde{p}_i \delta \tilde{q}_i - \tilde{H}_1 \delta t = c\,\Big(\sum_{i=1}^{n} p_i \delta q_i - H_1 \delta t\Big) - \delta F$$

Thus, (9) holds for any function H_1 and the corresponding function $\tilde{H}_1$.

Canonical transformations are sometimes also called *contact transformations*.

In the literature, only univalent canonical transformations are frequently considered, and many authors erroneously hold that these transformations exhaust all the transformations (1) that carry Hamiltonian systems again into Hamiltonian systems. These authors do not notice the arbitrary constant factor c which must figure in the general formula for an arbitrary canonical transformation.

---

* It does not follow from this however that a canonical transformation may be defined as a transformation that carries one given Hamiltonian system into some other Hamiltonian system.

Thus, for example, the arbitrary noncanonical transformation $\tilde{q}_i = \tilde{q}_i\,(q_k, p_k)$, $\tilde{p}_i = \tilde{p}_i\,(q_k, p_k)$ $(i = 1, \ldots, n)$ carries the Hamiltonian system with $H \equiv 0$ into the Hamiltonian system with $\tilde{H} \equiv 0$.

25. Free Canonical Transformations

Let us investigate in more detail the so-called free canonical transformations. These transformations are characterized, in addition, by the inequality

$$\frac{\partial(\tilde{q}_1, \ldots, \tilde{q}_n)}{\partial(p_1, \ldots, p_n)} \neq 0 \tag{1}$$

that ensures the independence of the quantities t, $q_1, \ldots, q_n$, $\tilde{q}_1, \ldots, \tilde{q}_n$ which can now be taken as the basic variables.* Indeed, this inequality enables us, using the first n equations (1), Sec. 24, to express the generalized momenta $p_1, \ldots, p_n$ in terms of $2n+1$ quantities t, q_i, $\tilde{q}_i$ $(i = 1, \ldots, n)$ and, consequently, to represent any function of the variables t, q_i, p_i $(i = 1, \ldots, n)$ as a function of the variables t, q_i, $\tilde{q}_i$ $(i = 1, \ldots, n)$. In this case it may be taken that the generating function is represented as a function S of these variables:

$$F(t, q_i, p_i) = S(t, q_i, \tilde{q}_i)$$

Then the basic identity (9) of the preceding section will be written as

$$\sum_{i=1}^{n} \tilde{p}_i \delta\tilde{q}_i - \tilde{H}\delta t = c\left(\sum_{i=1}^{n} p_i \delta q_i - H\delta t\right) - \delta S(t, q_i, \tilde{q}_i) \tag{2}$$

Equating the coefficients of δq_i, $\delta\tilde{q}_i$ and δt on the left and the right, we get the following formulas

$$\frac{\partial S}{\partial q_i} = cp_i, \quad \frac{\partial S}{\partial \tilde{q}_i} = -\tilde{p}_i \quad (i = 1, \ldots, n) \tag{3}$$

$$\tilde{H} = cH + \frac{\partial S}{\partial t} \tag{4}$$

Equations (3) define the canonical transformation under consideration. We shall show that they can be reduced to the form (1), Sec. 24.

The partial derivatives $\frac{\partial S}{\partial q_i}$ $(i = 1, \ldots, n)$ on the left-hand sides of the first n equations (3), as functions of the quantities $\tilde{q}_1, \ldots, \tilde{q}_n$ are independent because by virtue of (3) the relation**

$$\Omega\left(\frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}, q_1, \ldots, q_n\right) = 0$$

---

[*] In the case of a nonfree canonical transformation, the quantities t, q_i, $\tilde{q}_i$ $(i = 1, \ldots, n)$ are connected by certain relations.

[**] In this relation, the quantities $q_1, \ldots, q_n$ are taken as parameters.

would pass into the equality

$$\Omega(cp_1, \ldots, cp_n, q_1, \ldots, q_n) = 0$$

but this is only possible when $\Omega \equiv 0$, since the coordinates of a point of phase space q_i, p_i $(i = 1, \ldots, n)$ are independent quantities. From the independence of the derivatives $\frac{\partial S}{\partial q_i}$ $(i = 1, \ldots, n)$ considered as functions of the variables $\tilde{q}_1, \ldots, \tilde{q}_n$ it follows that the Jacobian of these functions is not identically zero. Thus, for the generating functions S of a free canonical transformation the determinant must be different from zero, i. e.:

$$\det\left(\frac{\partial^2 S}{\partial q_i \, \partial \tilde{q}_k}\right)_{i,\, k=1}^{n} \neq 0 \tag{5}$$

From the inequality (5) it follows that the first n equations (3) may be solved for $\tilde{q}_i$ $(i = 1, \ldots, n)$ and thus all the new phase variables $\tilde{q}_i$, $\tilde{p}_i$ $(i = 1, \ldots, n)$ may be expressed in terms of the old variables q_i, p_i $(i = 1, \ldots, n)$. The transformation of the form (1) thus obtained will be reversible, i.e. for it $\frac{\partial(\tilde{q}_1, \ldots, \tilde{p}_n)}{\partial(q_1, \ldots, p_n)} \neq 0$ since by virtue of inequality (5) the last n equations of (3) may be solved for q_i and, consequently, all the q_i, p_i may be expressed in terms of $\tilde{q}_i$, $\tilde{p}_i$ $(i = 1, \ldots, n)$. Thus, the equations (3) define a free canonical transformation with a given generating function $S(t, q_i, \tilde{q}_i)$ and with a given valence $c \neq 0$, while the formulas (4) establish a simple relation between the Hamiltonian functions H and $\tilde{H}$.

By running through the different generating functions S that satisfy the condition (5), and the different valences $c \neq 0$, we can obtain all the free canonical transformations* with the aid of the formulas (3).

For univalent $(c = 1)$ free canonical transformations, the formulas (3) and (4) take on a simpler form:

$$\frac{\partial S}{\partial q_i} = p_i, \quad \frac{\partial S}{\partial \tilde{q}_i} = -\tilde{p}_i \quad (i = 1, \ldots, n) \tag{6}$$

and

$$\tilde{H} = H + \frac{\partial S}{\partial t} \tag{7}$$

The last equation shows that after applying one and the same free univalent canonical transformation to various Hamiltonian systems, the difference between $\widetilde{H}$ and H will be one and the same $\left(\text{equal to } \frac{\partial S}{\partial t}\right)$ in all cases.

---

* For nonfree canonical transformations there exist defining formulas similar to (3). These formulas will be established in Sec. 29.

From (4) it follows that $\widetilde{H} = cH$ when and only when $\frac{\partial S}{\partial t} = 0$, i.e. when the generating function S does not depend explicitly on t. In this case, by virtue of (3), the time t will not enter explicitly into the formulas of the canonical transformation. Under such canonical transformations the function H does not change essentially, it is only multiplied by the constant c. For this reason, if we desire to obtain a new system with an essentially simpler Hamiltonian function, we must definitely take a free canonical transformation containing the time t as a parameter.

Example 1. On page 128 we considered three canonical transformations:

$$(1)\ \widetilde{q}_i = \alpha q_i, \quad \widetilde{p}_i = \beta p_i; \quad (2)\ \widetilde{q}_i = \alpha p_i, \quad \widetilde{p}_i = \beta q_i;$$

$$(3)\ \widetilde{q}_i = p_i \tan t, \quad \widetilde{p}_i = q_i \cot t \quad (i = 1, \ldots, n)$$

The transformations (2) and (3) are free. They have generating functions and valences, respectively, $S = -\beta \sum_{i=1}^{n} q_i\widetilde{q}_i$, $c = -\alpha\beta$, and $S = -\cot t \sum_{i=1}^{n} q_i\widetilde{q}_i$, $c = -1$. But the transformation (1) is not free. For it, $c = \alpha\beta$, $F \equiv 0$.

Example 2. Consider an arbitrary affine transformation of the phase plane (q, p) (with $n = 1$):

$$\left.\begin{aligned} \widetilde{q} &= \alpha q + \beta p, \\ \widetilde{p} &= \alpha_1 q + \beta_1 p \qquad (\alpha\beta_1 - \alpha_1\beta \neq 0) \end{aligned}\right\} \tag{8}$$

Put the right-hand sides of (8) into the basic identity (9) on page 130. Since the variable t does not enter into the right-hand sides of (8), we shall also seek F in the form of a function not dependent on the time explicitly: $F = F(q, p)$. Then the basic identity will take on the form

$$(\alpha_1 q + \beta_1 p)\ (\alpha\delta q + \beta\delta p) - cp\delta q = -\delta F$$

or

$$\delta\left(\frac{1}{2}\alpha\alpha_1 q^2 + \frac{1}{2}\beta\beta_1 p^2\right) + \alpha_1\beta q\delta p + (\alpha\beta_1 - c)\ p\delta q = -\delta F$$

The left-hand side of this equation will be a total differential provided that

$$c = \alpha\beta_1 - \alpha_1\beta$$

Thus, the transformation (8) is canonical with valence c, equal to the determinant of the transformation, and with the generating function

$$F = \frac{1}{2}\alpha\alpha_1 q^2 + \frac{1}{2}\beta\beta_1 p^2 + \alpha_1\beta q p$$

This transformation will be free if $\beta \neq 0$.

Example 3. The transformation $\widetilde{q} = \sqrt{q}\cos 2p$, $\widetilde{p} = \sqrt{q}\sin 2p$ is a free univalent canonical transformation with the generating function

$$S = \frac{1}{2} q \arccos \frac{\widetilde{q}}{\sqrt{q}} - \frac{1}{2}\widetilde{q}\sqrt{q - \widetilde{q}^2}$$

For a natural system, the coordinates $q_1, \ldots, q_n$ define the position of the system, and together with the momenta $p_1, \ldots, p_n$ they define the state of the system, that is, the positions and velocities of its points. This specificity of the coordinates is lost in a general-type canonical transformation. The quantities $\tilde{q}_1, \ldots, \tilde{q}_n$ no longer define the position of the system, and only together with the $\tilde{p}_1, \ldots, \tilde{p}_n$ do they define the state of the system. The variables $\tilde{q}_1, \ldots, \tilde{q}_n$ will as before define the position of the system only in the particular case of a point canonical transformation for which the functions $\tilde{q}_i(t, q_k, p_k)$ actually do not contain the momenta:

$$\tilde{q}_i = \tilde{q}_i(t, q_k) \quad (i = 1, \ldots, n)$$

Note that subsequently the transformation of an arbitrary Hamiltonian system into a system with the function H of simple structure may be effected with the aid of a free canonical transformation. But a free canonical transformation is not a point transformation. Thus, nonpoint canonical transformations play an essential role in the theory of Hamiltonian systems.

26. The Hamilton-Jacobi Equation

The theory of canonical transformations leads us directly to the Hamilton-Jacobi equation.

Let there be given a holonomic system whose motion obeys the canonical equations of Hamilton:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad (i = 1, \ldots, n) \tag{1}$$

We shall try to determine a free univalent canonical transformation such that in the transformed Hamiltonian system

$$\frac{d\tilde{q}_i}{dt} = \frac{\partial \tilde{H}}{\partial \tilde{p}_i}, \quad \frac{d\tilde{p}_i}{dt} = -\frac{\partial \tilde{H}}{\partial \tilde{q}_i} \quad (i = 1, \ldots, n) \tag{2}$$

the function $\tilde{H}$ will be identically zero:

$$\tilde{H} \equiv 0 \tag{3}$$

Then the system (2) is integrated directly:

$$\tilde{q}_i = \alpha_i, \quad \tilde{p}_i = \beta_i \tag{4}$$

where α_i and β_i are $2n$ arbitrary constants. Knowing the canonical transformation, i.e. the relation between q_i, p_i $(i = 1, \ldots, n)$ and $\tilde{q}_i$, $\tilde{p}_i$ $(i = 1, \ldots, n)$, we can express all the q_i and p_i as functions of the time t and of the $2n$ arbitrary constants α_k, β_k $(k =$

$= 1, \ldots, n$), that is we can find the final equations of motion of the given holonomic system completely [all the solutions of the system (1)].

How are we to determine the canonical transformation that we need? To do this, by virtue of formula (7) of the preceding section, it is necessary and sufficient that the equation

$$\frac{\partial S}{\partial t} + H(t, q_i, p_i) = 0 \tag{5}$$

hold for the generating function $S(t, q_i, \tilde{q}_i)$ of the desired canonical transformation. This, together with the formulas (6) of the same section, yields

$$\frac{\partial S}{\partial t} + H\left(t, q_i, \frac{\partial S}{\partial q_i}\right) = 0 \tag{6}$$

The resulting partial differential equation (6) is called the *Hamilton-Jacobi equation.* Thus the generating function $S(t, q_i, \tilde{q}_i)$ with basic variables t and q_i (the $\tilde{q}_i$ are regarded as parameters) satisfies the partial differential equation of Hamilton-Jacobi. Besides the Hamilton-Jacobi equation, the condition

$$\det\left(\frac{\partial^2 S}{\partial q_i\, \partial \tilde{q}_k}\right)_{i,\, k=1}^{n} \neq 0 \tag{7}$$

must hold for the generating function $S(t, q_i, \tilde{q}_i)$.

As soon as the generating function $S(t, q_i, \tilde{q}_i)$ is found, the formulas

$$\frac{\partial S}{\partial q_i} = p_i, \quad \frac{\partial S}{\partial \tilde{q}_i} = -\tilde{p}_i \quad (i = 1, \ldots, n)$$

will define the sought-for free canonical transformation. If in these formulas we replace the $\tilde{q}_i$ by α_i and the $\tilde{p}_i$ by β_i, we will get the equations of motion of the given holonomic system in closed form. This whole process is more conveniently described if from the very start the $\tilde{q}_i$ in S are replaced by α_i $(i = 1, \ldots, n)$.

We introduce a definition. The solution $S(t, q_i, \alpha_i)$ of the partial differential equation of Hamilton-Jacobi containing n arbitrary constants $\alpha_1, \ldots, \alpha_n$ is called the *complete integral* of this equation if the condition

$$\det\left(\frac{\partial^2 S}{\partial q_i\, \partial \alpha_k}\right)_{i,\, k=1}^{n} \neq 0 \tag{8}$$

is fulfilled.

Now let us formulate the theorem we have proved.

Jacobi's theorem. *If $S(t, q_i, \alpha_i)$ is some complete integral of the Hamilton-Jacobi equation* (6), *then the final equations of motion of* *a holonomic system with the given function H may be written in the form**

$$\frac{\partial S}{\partial q_i}=p_i,\ \frac{\partial S}{\partial \alpha_i}=\beta_i\quad (i=1,\ \ldots,\ n)\tag{9}$$

where α_i and β_i are arbitrary constants ($i = 1, \ldots, n$).

Thus, a knowledge of the complete integral of the partial differential equation (6) relieves us of the necessity of integrating the system of ordinary differential equations (1). The problem of integrating this system is replaced by an equivalent problem of seeking the complete integral of the partial differential equation of Hamilton-Jacobi.

Note. The general solution of the partial differential equation depends on several arbitrary functions. For this reason a complete integral of the Hamilton-Jacobi equation is by no means a general solution. A complete integral embraces only a small number of solutions when compared with a general solution. It is nevertheless possible, on the basis of a complete integral, to restore the initial equation (whence the name "complete" integral). Indeed, differentiating the complete integral, we get

$$\frac{\partial S}{\partial q_i}=f_i(t,\ q_k,\ \alpha_k)\quad (i=1,\ \ldots,\ n)\tag{10}$$

$$\frac{\partial S}{\partial t}=f_0(t,\ q_k,\ \alpha_k)\tag{11}$$

If the complete integral $S(t,\ q_k,\ \alpha_k)$ is known, then the functions $f_i(t,\ q_k,\ \alpha_k)$ ($i=0,\ 1,\ \ldots,\ n$) are also known. From relation (10) it is possible to express every α_k in terms of the partial derivatives $\frac{\partial S}{\partial q_i}$, t and q_i since, by virtue of condition (8),

$$\frac{\partial (f_1,\ \ldots,\ f_n)}{\partial (\alpha_1,\ \ldots,\ \alpha_n)}=\det\left(\frac{\partial^2 S}{\partial q_i\,\partial \alpha_k}\right)_{i,\ k=1}^{n}\neq 0\tag{12}$$

Putting the expressions obtained for α_k into (11) we get the initial partial differential equation (6).**

As an instance of a complete integral of the Hamilton-Jacobi equation, we consider what is called *Hamilton's principal function.* To do this, we return to formula (7) on page 99 and to Fig. 33 on page 100. We now consider only the special case when $t_0(\alpha) = \text{const} = t_0$, i.e. we assume that the contour C_0 consists of the initial states of the system when $t = t_0$. Besides, in place of t_1, q_i^1, p_i^1, H_1 we shall write simply t, q_i, p_i, H. Then, if W is the action along a straight-line path (i.e. along the generatrix of the tube) from the initial point ($t = t_0$) to the terminal point corresponding to the given value of t, we get

---

* Here, in place of the arbitrary constants $-\beta_i$ we write simply β_i. By virtue ot the condition (8), the last n equations (9) may be solved for q_i and we can express $q_1, \ldots, q_n$ as functions of t and the $2n$ arbitrary constants α_i, β_i.

** It may be considered that the complete integral S also contains an ($n+1$)st additive arbitrary constant α_{n+1}, since only the derivatives of S enter into equation (6) and not the function S itself.

$$\delta W = \sum_{i=1}^{n} p_i \delta q_i - H\delta t - \sum_{i=1}^{n} p_i^0 \delta q_i^0 \tag{13}$$

If we take advantage of the final equations of motion

$$\left.\begin{aligned} q_i &= \varphi_i (t, q_k^0, p_k^0) \qquad (14)\\ p_i &= \psi_i (t, q_k^0, p_k^0) \qquad (15)\end{aligned}\right\} \quad (i = 1, \ldots, n)$$

and in place of $q_i(t)$ put their values (14) into the expression for the action

$$W = \int_{t_0}^{t} L(t, q_i, \dot{q}_i)\, dt$$

then W will be a function of t, q_i^0, p_i^0 ($i = 1, \ldots, n$). Hamilton proposed using the final equations of motion (14) to express p_k^0 in terms of t, q_i^0 and q_i and thus to represent the action in the form

$$W = W(t, q_i, q_i^0) \tag{16}$$

The action W given in the form (16), i.e. in the form of a function of the initial coordinates, the terminal coordinates and the terminal instant of time t, is called *Hamilton's principal function.* Considering that in (13) W is Hamilton's principal function, we get (on the basis of this equation)

$$\frac{\partial W}{\partial q_i} = p_i, \quad \frac{\partial W}{\partial q_i^0} = -p_i^0 \quad (i = 1, \ldots, n) \tag{17}$$

and

$$\frac{\partial W}{\partial t} = -H(t, q_i, p_i) \tag{18}$$

From the equations (17) and (18) it follows that Hamilton's principal function satisfies the Hamilton-Jacobi equation

$$\frac{\partial W}{\partial t} + H\left(t, q_i, \frac{\partial W}{\partial q_i}\right) = 0 \tag{19}$$

and the relations (17) are the final equations of motion containing $2n$ arbitrary constants q_i^0, p_i^0 ($i = 1, \ldots, n$).

Thus, Hamilton demonstrated how the final equations of motion are written by means of a complete integral of the equation (19). However, in Hamilton's case this complete integral was not arbitrary; and the arbitrary constants in it were the initial values of the

q_i^0 and the p_i^0. The result was a vicious circle: to write the final equations of motion (17) one needs Hamilton's principal function, and to construct this function, as has already been demonstrated above, one needs to know the final equations of motion.

Jacobi's contribution consists in the fact that he continued Hamilton's investigations and broke the vicious circle. He demonstrated that the final equations of motion might be written in the form (9) by means of an *arbitrary* complete integral $S\ (t, q_i, \alpha_i)$ of the Hamilton-Jacobi equation.

Let us return to the identity (13) and compare it with the identity (2) on page 132. It will be seen from this comparison that the formulas (14) and (15), which are the final equations of motion and which express the Hamiltonian coordinates q_i, p_i of the state of the system at time t in terms of the initial coordinates $q_1^0, \ldots, p_n^0$, may be regarded as a free univalent canonical transformation from the variables q_i^0, p_i^0 $(i = 1, \ldots, n)$ to the variables q_k, p_k $(k = 1, \ldots, n)$; $-W$ is the generating function of this canonical transformation, and W is Hamilton's principal function.*

Thus, a transformation of phase space carried out by means of the motions of any Hamiltonian system is canonical (and also free and univalent).

Example. Form the Hamiltonian function for the inertial motion of a free particle. In this case (assume $t_0 = 0$)

$$x = x_0 + \dot{x}_0 t, \quad y = y_0 + \dot{y}_0 t, \quad z = z_0 + \dot{z}_0 t$$

and so

$$W = \frac{m}{2} \int_0^t (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)\, dt = \frac{mt}{2} (\dot{x}_0^2 + \dot{y}_0^2 + \dot{z}_0^2) = \frac{m}{2t} [(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2]$$

If we proceed from Hamilton's principal function that we found

$$W = \frac{m}{2t} [(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2]$$

then the equations of motion are obtainable from the formulas (17), which in the given instance appear as

$$p_x = \frac{\partial W}{\partial x} = m \frac{x - x_0}{t}; \quad -p_x^0 = \frac{\partial W}{\partial x_0} = -m \frac{x - x_0}{t}$$

$$p_y = \frac{\partial W}{\partial y} = m \frac{y - y_0}{t}; \quad -p_y^0 = \frac{\partial W}{\partial y_0} = -m \frac{y - y_0}{t}$$

$$p_z = \frac{\partial W}{\partial z} = m \frac{z - z_0}{t}; \quad -p_z^0 = \frac{\partial W}{\partial z_0} = -m \frac{z - z_0}{t}$$

---

* For the inverse canonical transformation from the variables q_i, p_i to the variables q_i^0, p_i^0, Hamilton's principal function W will be the generating function.

Suppose we have a generalized-conservative system $\left(\frac{\partial H}{\partial t}=0\right)$. Then the Hamilton-Jacobi equation has the form

$$\frac{\partial S}{\partial t}+H\left(q_i, \frac{\partial S}{\partial q_i}\right)=0 \tag{20}$$

and its complete solution may be found in the form

$$S=-ht+V(q_1, \ldots, q_n; \alpha_1, \ldots, \alpha_{n-1}, h) \tag{21}$$

where h and $\alpha_1, \ldots, \alpha_{n-1}$ are arbitrary constants.

Putting this expression for S in equation (20), we get the following equation for determining the function V:

$$H\left(q_i, \frac{\partial V}{\partial q_i}\right)=h \tag{22}$$

Finding the complete integral of this equation, i.e. the solution $V(q_i; \alpha_1, \ldots, \alpha_{n-1}, h)$ for which the inequality

$$\det\left(\frac{\partial^2 V}{\partial q_i\, \partial \alpha_k}\right)_{i,\ k=1}^{n} \neq 0 \quad (\alpha_n \equiv h) \tag{23}$$

holds, we obtain with the aid of formulas (9) and (21) the following final equations of motion of a generalized-conservative system:

$$\frac{\partial V}{\partial q_i}=p_i \quad (i=1, \ldots, n) \tag{24'}$$

$$\begin{aligned}\frac{\partial V}{\partial \alpha_j}&=\beta_j \quad (j=1, \ldots, n-1),\\ \frac{\partial V}{\partial h}&=t+\gamma\end{aligned} \tag{24''}$$

where α_j, β_j $(j=1, \ldots, n-1)$, h and γ are arbitrary constants.

By virtue of the condition (23) the coordinates $q_1, \ldots, q_n$ may be determined from the equation (24″) as functions of t and the $2n$ arbitrary constants α_j, β_j, h and γ. Putting the expressions obtained into equations (24′), we get similar expressions for the generalized momenta p_i $(i=1, \ldots, n)$.*

In the case of a generalized-conservative system we replaced equation (6) by equation (20), in which the number of independent variables is less by one. A similar reduction in the number of independent variables in the partial differential equation may be also accomplished when one of the coordinates is cyclic (ignorable).

Let us consider at once the general case when several coordinates $q_{m+1}, \ldots, q_n$ are cyclic. Then $H=H(t, q_1, \ldots, q_m, p_1, \ldots, p_n)$, and the complete solution of the Hamilton-Jacobi equation may be sought in the form

---

* Using the first $n-1$ equations of (24″) it is possible to express $n-1$ coordinates in terms of the remaining nth coordinate and $2n-1$ arbitrary constants. We then obtain the equations of trajectories in coordinate space. The last equation of (24″) connects the coordinates with the time variable t.

$$S=\sum_{\mu=m+1}^{n}\alpha_\mu q_\mu+S_0(t,\ q_1,\ \ldots,\ q_m,\ \alpha_1,\ \ldots,\ \alpha_n) \tag{25}$$

For the function S_0 we get the equation

$$\frac{\partial S_0}{\partial t}+H\left(t,\ q_1,\ \ldots,\ q_m,\ \frac{\partial S_0}{\partial q_1},\ \ldots,\ \frac{\partial S_0}{\partial q_m},\ \alpha_{m+1},\ \ldots,\ \alpha_n\right)=0 \tag{26}$$

If in a generalized-conservative system the coordinates $q_{m+1},\ \ldots,\ q_n$ are cyclic, then the function S is sought in the form

$$S=-ht+\sum_{\mu=m+1}^{n}\alpha_\mu q_\mu+V_0(q_1,\ \ldots,\ q_m,\ \alpha_1,\ \ldots,\ \alpha_n,\ h) \tag{27}$$

where the function V_0 is determined from the equation

$$H\left(q_1,\ \ldots,\ q_m,\ \frac{\partial V_0}{\partial q_1},\ \ldots,\ \frac{\partial V_0}{\partial q_m},\ \alpha_{m+1},\ \ldots,\ \alpha_n\right)=h \tag{28}$$

27. Method of Separation of Variables. Examples

We have shown that integrating a system of canonical equations reduces to finding a complete solution of the Hamilton-Jacobi equation. The interest here is not only theoretical. It turns out that many problems of dynamics, including problems of interest in theoretical physics, may be conveniently solved in this way.

We shall now examine the method of separation of variables for finding a complete solution to the Hamilton-Jacobi equation. This method is applied in cases when the Hamiltonian function H of a generalized-conservative system has a special structure.

1°. Let

$$H=G[f_1(q_1,\ p_1),\ \ldots,\ f_n(q_n,\ p_n)] \tag{1}$$

Here, the variables in the expression for the function H are separated, i.e. only one pair of conjugate variables q_i, p_i enters into each function f_i.

Equation (22) of the preceding section will now be written as

$$G\left[f_1\left(q_1,\ \frac{\partial V}{\partial q_1}\right),\ \ldots,\ f_n\left(q_n,\ \frac{\partial V}{\partial q_n}\right)\right]=h \tag{2}$$

Put

$$f_i\left(q_i,\ \frac{\partial V}{\partial q_i}\right)=\alpha_i \quad (i=1,\ \ldots,\ n) \tag{3}$$

where $\alpha_1, \ldots, \alpha_n$ are arbitrary constants. Then, by formula (1), the constant h may be expressed in terms of the constants $\alpha_1, \ldots, \alpha_n$ as follows:

$$h = G(\alpha_1, \ldots, \alpha_n) \tag{4}$$

Solving (3) for $\frac{\partial V}{\partial q_i}$, we find*

$$\frac{\partial V}{\partial q_i} = F_i(q_i, \alpha_i) \quad (i = 1, \ldots, n),$$

$$V = \sum_{i=1}^{n} \int F_i(q_i, \alpha_i)\, dq_i \tag{5}$$

and

$$S = -G(\alpha_1, \ldots, \alpha_n)\, t + \sum_{i=1}^{n} \int F_i(q_i, \alpha_i)\, dq_i \tag{6}$$

In this case

$$\frac{\partial^2 S}{\partial q_i\, \partial \alpha_i} = \frac{\partial F_i}{\partial \alpha_i}, \quad \frac{\partial^2 S}{\partial q_i\, \partial \alpha_k} = 0 \quad \text{for} \quad i \neq k \quad (i, k = 1, \ldots, n)$$

and the basic condition

$$\det \left(\frac{\partial^2 S}{\partial q_i\, \partial \alpha_k} \right)_{i,\, k=1}^{n} \neq 0 \tag{7}$$

reduces to the inequality

$$\prod_{i=1}^{n} \frac{\partial F_i}{\partial \alpha_i} \neq 0$$

Since the relation

$$f_i(q_i, p_i) = \alpha_i \tag{8}$$

is equivalent to the equation

$$p_i = F_i(q_i, \alpha_i) \tag{9}$$

it follows that

$$\frac{\partial F_i}{\partial \alpha_i} = \left(\frac{\partial f_i}{\partial p_i} \right)^{-1} \neq 0 \quad (i = 1, \ldots, n) \tag{10}$$

---

* We assume that each function $f_i(q_i, p_i)$ actually contains the momentum p_i, i.e. that $\frac{\partial f_i}{\partial p_i} \neq 0$. In this case equation (3) may be solved for $p_i = \frac{\partial V}{\partial q_i}$; each function F_i is a function of the two variables q_i and α_i; $i = 1, \ldots, n$.

and condition (7) always holds. Therefore formula (6) defines a complete integral of the Hamilton-Jacobi equation.

In the given case the final equations of motion

$$\frac{\partial S}{\partial q_i}=p_i,\quad \frac{\partial S}{\partial \alpha_i}=\beta_i \quad (i=1,\ \ldots,\ n) \tag{11}$$

will be written as*

$$\left.\begin{aligned} -\frac{\partial G}{\partial \alpha_i}t+\int\frac{dq_i}{\left(\frac{\partial f_i}{\partial p_i}\right)_{p_i=F_i(q_i,\ \alpha_i)}}&=\beta_i, \\ p_i=F_i(q_i,\ \alpha_i) \quad (i=1,\ \ldots,\ n)& \end{aligned}\right\} \tag{12}$$

Thus, the final equations of motion are found by quadratures.

Example 1. Consider an oscillator with one degree of freedom. Here

$$H=\frac{p^2}{2m}+\frac{cq^2}{2}$$

and, thus,

$$f(p,\ q)=\frac{p^2}{2m}+\frac{cq^2}{2}$$

and equation (2) is of the form

$$\frac{1}{2m}\left(\frac{dV}{dq}\right)^2+\frac{cq^2}{2}=h$$

Put $h=\alpha$; then

$$S=-\alpha t+\int\sqrt{2m\alpha-mcq^2}\,dq$$

and from the final equations of motion (11) we find the expression for the momentum.

$$p=\sqrt{2m\alpha-mcq^2} \tag{13}$$

and the equation of motion itself will be written as

$$-t+\frac{1}{\omega}\int\frac{dq}{\sqrt{A^2-q^2}}=\beta \tag{14}$$

---

* In (5) and (6) by the integral $\int F_i(q_i,\ \alpha_i)\,dq_i$ we mean the integral $\int\limits_{\gamma_i}^{q_i} F_i(q_i,\ \alpha_i)\,dq_i$ where the constant γ_i is fixed and is independent of the values of the arbitrary constants α_k; then

$$\frac{\partial}{\partial \alpha_i}\sum_{k=1}^{n}\int F_k(q_k,\ \alpha_k)\,dq_k=\frac{\partial}{\partial \alpha_i}\int F_i(q_i,\ \alpha_i)\,dq_i=\int\frac{\partial F_i}{\partial \alpha_i}\,dq_i$$

We then make use of (10).

where

$$A^2 = \frac{2\alpha}{c}, \quad \omega^2 = \frac{c}{m}$$

Noting that the integral in (14) is equal to arc sin $\frac{q}{A}$, we get the equation of motion in the following form:

$$q = A \sin \omega (t + \beta) \tag{15}$$

2°. Let

$$H = g_n \{\dots g_3 \{g_2 [g_1 (q_1, p_1), q_2, p_2], q_3, p_3\} \dots q_n, p_n\} \tag{16}$$

Then the equation determining V will be written thus

$$g_n \left\{\dots g_3 \left\{ g_2 \left[g_1 \left(q_1, \frac{\partial V}{\partial q_1} \right), q_2, \frac{\partial V}{\partial q_2} \right], q_3, \frac{\partial V}{\partial q_3} \right\} \dots q_n, \frac{\partial V}{\partial q_n} \right\} = h$$

We introduce the arbitrary variables $\alpha_1, \dots, \alpha_{n-1}$ and $\alpha_n = h$ and, successively, put

$$\left.\begin{aligned} g_1 \left(q_1, \frac{\partial V}{\partial q_1} \right) &= \alpha_1, \\ g_2 \left(\alpha_1, q_2, \frac{\partial V}{\partial q_2} \right) &= \alpha_2, \\ \dots\dots\dots\dots \\ g_n \left(\alpha_{n-1}, q_n, \frac{\partial V}{\partial q_n} \right) &= \alpha_n \end{aligned}\right\} \tag{17}$$

Determining from this the partial derivatives, we find*

$$\begin{aligned} \frac{\partial V}{\partial q_1} &= G_1 (q_1, \alpha_1), \\ \frac{\partial V}{\partial q_2} &= G_2 (q_2, \alpha_1, \alpha_2), \\ \dots\dots\dots \\ \frac{\partial V}{\partial q_n} &= G_n (q_n, \alpha_{n-1}, \alpha_n) \end{aligned}$$

Whence

$$V = \sum_{i=1}^{n} \int G_i (q_i, \alpha_{i-1}, \alpha_i) \, dq_i \, ^{**} \tag{18}$$

$$S = -\alpha_n t + \sum_{i=1}^{n} \int G_i (q_i, \alpha_{i-1}, \alpha_i) \, dq_i \tag{19}$$

---

* We assume that $\frac{\partial g_i}{\partial p_i} \neq 0$, i.e. that p_i actually enters into the function $g_i (q_i, p_i)$.

** Here and henceforward in this section, the indefinite integral is understood to mean (as in Item 1°) the definite integral with variable upper limit and fixed constant lower limit independent of $\alpha_1, \dots, \alpha_n$.

Here

$$\frac{\partial^2 S}{\partial q_i \partial \alpha_i} = \frac{\partial G_i}{\partial \alpha_i}, \quad \frac{\partial^2 S}{\partial q_i \partial \alpha_k} = 0$$

$$\text{for } i < k \, (i, \; k = 1, \ldots, n)$$

And so the condition (7) reduces to the inequality

$$\prod_{i=1}^{n} \frac{\partial G_i}{\partial \alpha_i} \neq 0$$

which always holds since the equation

$$g_i(\alpha_{i-1}, \; q_i, \; p_i) = \alpha_i \tag{20}$$

is equivalent to the equation

$$p_i = G_i(q_i, \; \alpha_{i-1}, \; \alpha_i) \tag{21}$$

and for this reason

$$\frac{\partial G_i}{\partial \alpha_i} = \left(\frac{\partial g_i}{\partial p_i} \right)^{-1}_{p_i = G_i(q_i, \, \alpha_{i-1}, \, \alpha_i)} \neq 0 \quad (i = 1, \ldots, n) \tag{22}$$

For what is to follow we need the expressions for the derivatives $\frac{\partial G_i}{\partial \alpha_{i-1}}$, which are found from equations (20) and (21), namely:

$$\frac{\partial G_i}{\partial \alpha_{i-1}} = - \left(\frac{\dfrac{\partial g_i}{\partial \alpha_{i-1}}}{\dfrac{\partial g_i}{\partial p_i}} \right)_{p_i = G_i(q_i, \, \alpha_{i-1}, \, \alpha_i)} \tag{23}$$

Putting expression (19) into the final equations of motion (11) and taking into account the formulas (22) and (23), we finally get

$$\left. \begin{array}{c} \displaystyle \int \frac{dq_i}{\left(\dfrac{\partial g_i}{\partial p_i} \right)_{p_i = G_i \, (q_i, \, \alpha_{i-1}, \, \alpha_i)}} - \\ \displaystyle - \int \left(\frac{\dfrac{\partial g_{i+1}}{\partial \alpha_i}}{\dfrac{\partial g_{i+1}}{\partial p_{i+1}}} \right)_{p_{i+1} = G_{i+1}(q_{i+1}, \, \alpha_i, \, \alpha_{i+1})} dq_{i+1} = \beta_i \\ (i = 1, \ldots, n-1), \\ \displaystyle -t + \int \frac{dq_n}{\left(\dfrac{\partial g_n}{\partial p_n} \right)_{p_n = G_n(q_n, \, \alpha_{n-1}, \, \alpha_n)}} = \beta_n \end{array} \right\} \tag{24}$$

and

$$p_i = G_i(q_i,\ \alpha_{i-1},\ \alpha_i) \quad (i = 1, \ldots, n) \tag{25}$$

Here, the first $n-1$ equations of (24) are equations for a family of trajectories in coordinate space; these equations contain $2n-1$ arbitrary constants $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_{n-1}$. The last equation of (24) contains a new arbitrary constant β_n and connects the coordinates q_i with the time variable t.

After substituting into equations (25) the functions $q_i\,(t, \alpha_k, \beta_k)$ $(i = 1, \ldots, n)$ found from (24), equations (25) determine the momenta p_i $(i = 1, \ldots, n)$ as functions of t and all arbitrary constants α_i, β_i $(i = 1, \ldots, n)$.

Example 2. Consider Keplerian motion in which a particle of mass m is attracted to a centre with a force inversely proportional to the square of the distance from the centre.

In this case, in spherical coordinates,

$$T = \frac{m}{2}\,(\dot{r}^2 + r^2\dot{\Theta}^2 + r^2\sin^2\theta\,\dot{\psi}^2)$$

$$\Pi = -\frac{\gamma}{r} \quad (\gamma > 0)$$

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\psi^2}{r^2\sin^2\theta}\right) - \frac{\gamma}{r}$$

and the equation that determines V is of the form

$$\frac{1}{2m}\left\{\left(\frac{\partial V}{\partial r}\right)^2 + \frac{1}{r^2}\left[\left(\frac{\partial V}{\partial\theta}\right)^2 + \right.\right.$$

$$\left.\left. + \frac{1}{\sin^2\theta}\left(\frac{\partial V}{\partial\psi}\right)^2\right]\right\} - \frac{\gamma}{r} = h$$

Put

$$g_1 \equiv p_\psi = \alpha_1, \quad g_2 \equiv p_\theta^2 + \frac{\alpha_1^2}{\sin^2\theta} = \alpha_2,$$

$$g_3 \equiv \frac{1}{2m}\left(p_r^2 + \frac{\alpha_2}{r^2}\right) - \frac{\gamma}{r} = \alpha_3 = h$$

Then, by formulas (24), we find

$$\psi - \alpha_1 \int \frac{d\theta}{\sin^2\theta\sqrt{\alpha_2 - \dfrac{\alpha_1^2}{\sin^2\theta}}} = \beta_1, \qquad \text{(26a}$$

$$\int \frac{d\theta}{2\sqrt{\alpha_2 - \dfrac{\alpha_1^2}{\sin^2\theta}}} -$$

$$-\int \frac{\frac{1}{r^2}\,dr}{2\sqrt{2m\alpha_3+\frac{2m\gamma}{r}-\frac{\alpha_2}{r^2}}}=\beta_2, \tag{26b}$$

$$-t+\int \frac{m\,dr}{\sqrt{2m\alpha_3+\frac{2m\gamma}{r}-\frac{\alpha_2}{r^2}}}=\beta_3 \tag{26c}$$

We have obtained the final equations for Keplerian motion.

In studying this motion, it may be taken, without loss of generality, that the initial velocity lies in the plane of the meridian $\psi=\text{const}$. Then at the initial time, $\frac{\partial\psi}{\partial\theta}=0$ and, hence, by formula (26a),

$$\alpha_1=0 \tag{27}$$

and $\psi=\beta_1=\text{const}$, i.e. the motion is plane motion. Differentiating the equations (26b) and (26c) termwise, we find that the sector velocity* is

$$\frac{1}{2}r^2\frac{d\theta}{dt}=\frac{\sqrt{\alpha_2}}{2m}=\text{const}$$

which means that the motion in the plane $\psi=\text{const}$ occurs in accordance with the law of areas.

Finally, to determine the trajectory we put $\frac{1}{r}=x$ and from (26b), taking (27) into consideration, we find

$$\int \frac{dx}{\sqrt{c+2kx-x^2}}=\beta-\theta$$

where

$$c=\frac{2m\alpha_3}{\alpha_2}, \quad k=\frac{m\gamma}{\alpha_2},$$

$$\beta=2\sqrt{\alpha_2}\,\beta_2 \tag{28}$$

Computing the integral we get

$$\arccos\frac{x-k}{\sqrt{k^2+c}}=0-\beta$$

and, hence,

$$x=k+\sqrt{k^2+c}\cos(\theta-\beta)$$

Finally, recalling that $x=\frac{1}{r}$, we have for the trajectory an equation of a conic section, in one focus of which is the centre of attraction:

$$r=\frac{p}{1+e\cos(\theta-\beta)} \tag{29}$$

---

* That is, the time derivative of the area described by a radius vector drawn from the centre.

where the parameter and eccentricity of the conic section are determined from the equations

$$p=\frac{1}{k}=\frac{\alpha_2}{m\gamma}, \quad e=\sqrt{1+\frac{c}{k^2}}=\sqrt{1+\frac{2\alpha_2\alpha_3}{m\gamma^2}} \tag{30}$$

Let a point describe a closed orbit (the motion of a planet attracted by the sun). Then this orbit is an ellipse at one focus of which is the centre of attraction (sun).

Denoting by F and a, b $(a>b)$ the area and the semi-axes of the ellipse, we find $\left(\text{since it is known that } p=\frac{b^2}{a}\right)$ that

$$\frac{F^2}{a^3}=\frac{\pi^2a^2b^2}{a^3}=\pi^2p$$

Let τ be the period (of revolution), $\tau=\frac{2mF}{\sqrt{\alpha_2}}$. Then on the basis of (30)

$$\frac{1}{4}\frac{\tau^2}{a^3}=\frac{m^2}{\alpha_2}\frac{F^2}{a^3}=\frac{\pi^2m^2p}{\alpha_2}=\frac{\pi^2m}{\gamma}=\frac{\pi^2}{\gamma_1}\left(\gamma_1=\frac{\gamma}{m}\right) \tag{31}$$

where by Newton's law of gravitation, γ_1 depends solely on the centre of attraction. We have obtained Kepler's three laws of planetary motion about the sun: (1) the planets are in motion with constant sector velocity in plane orbits; (2) these orbits are ellipses in one focus of which lies the sun; (3) the ratio of the squares of the orbital periods to the cubes of the major axes of the orbits is the same for all planets.

3°. By way of further illustrating the application of the method of separation of variables, consider the case when

$$H=\frac{f_1(q_1, p_1)+\ldots+f_n(q_n, p_n)}{g_1(q_1, p_1)+\ldots+g_n(q_n, p_n)} \tag{32}$$

Then the basic differential equation

$$H\left(q_1, \ldots, q_n, \frac{\partial V}{\partial q_1}, \ldots, \frac{\partial V}{\partial q_n}\right)=h$$

may be written as

$$\sum_{i=1}^{n}\left[f_i\left(q_i, \frac{\partial V}{\partial q_i}\right)-hg_i\left(q_i, \frac{\partial V}{\partial q_i}\right)\right]=0 \tag{33}$$

Put

$$f_i\left(q_i, \frac{\partial V}{\partial q_i}\right)-hg_i\left(q_i, \frac{\partial V}{\partial q_i}\right)=\alpha_i \quad (i=1, \ldots, n) \tag{34}$$

where $\alpha_1, \ldots, \alpha_{n-1}$ are arbitrary constants, and the constant α_n, by virtue of equation (33) and equalities (34), is expressed in terms of $\alpha_1, \ldots, \alpha_{n-1}$, namely

$$\alpha_n=-\alpha_1-\ldots-\alpha_{n-1} \tag{35}$$

Solve (34) for the derivatives $\frac{\partial V}{\partial q_i}$:

$$\frac{\partial V}{\partial q_i}=F_i(q_i, \alpha_i, h) \quad (i=1, \ldots, n) \tag{36}$$

Take the solution of (33) in the form

$$V=\sum_{i=1}^{n}\int F_i(q_i,\ \alpha_i,\ h)\,dq_i \tag{37}$$

Therefore a complete integral of the Hamilton-Jacobi equation has the form

$$S=-ht+\sum_{i=1}^{n}\int F_i(q_i,\ \alpha_i,\ h)\,dq_i \tag{38}$$

Then the final equations of motion (11) are obtained by means of quadratures:

$$\left.\begin{array}{c}
\displaystyle\int\frac{\partial F_j}{\partial\alpha_j}\,dq_j-\int\frac{\partial F_n}{\partial\alpha_n}\,dq_n=\beta_j\quad(j=1,\ \ldots,\ n-1),\\
\displaystyle\sum_{i=1}^{n}\int\frac{\partial F_i}{\partial h}\,dq_i=t+\beta_n,\\
p_i=F_i(q_i,\ \alpha_i,\ h)\quad(i=1,\ \ldots,\ n)\\
(\alpha_1+\alpha_2+\ldots+\alpha_n=0)
\end{array}\right\} \tag{39}$$

Solving the equation $f_i(q_i,\ p_i)-hg_i(q_i,\ p_i)=\alpha_i$ for p_i, we get $p_i=F_i(q_i,\ \alpha_i,\ h)$ $(i=1,\ \ldots,\ n)$. It is assumed that the following condition of solvability holds: $\frac{\partial f_i}{\partial p_i}-h\frac{\partial g_i}{\partial p_i}\neq 0$. In this case the derivatives of the implicit function F_i are

$$\frac{\partial F_i}{\partial\alpha_i}=\left[\frac{\partial f_i}{\partial p_i}-h\frac{\partial g_i}{\partial p_i}\right]^{-1}_{p_i=F_i(q_i,\ \alpha_i,\ h)}$$

and

$$\frac{\partial F_i}{\partial h}=\left\{\left[\frac{\partial f_i}{\partial p_i}-h\frac{\partial g_i}{\partial p_i}\right]^{-1}g_i(q_i,\ p_i)\right\}_{p_i=F_i(q_i,\ \alpha_i,\ h)}\qquad(i=1,\ \ldots,\ n)$$

In final form, the equations (39) are

$$\left.\begin{array}{c}
\displaystyle\int\left[\frac{\partial f_j}{\partial p_j}-h\frac{\partial g_j}{\partial p_j}\right]^{-1}_{p_j=F_j(q_j,\ \alpha_j,\ h)}dq_j-\\
\displaystyle-\int\left[\frac{\partial f_n}{\partial p_n}-h\frac{\partial g_n}{\partial p_n}\right]_{p_n=F_n(q_n,\ \alpha_n,\ h)}dq_n=\beta_j\quad(j=1,\ \ldots,\ n-1),\\
\displaystyle\sum_{i=1}^{n}\int\left\{\left[\frac{\partial f_i}{\partial p_i}-h\frac{\partial g_i}{\partial p_i}\right]^{-1}g_i(q_i,\ p_i)\right\}_{p_i=F_i(q_i,\ \alpha_i,\ h)}dq_i=t+\beta_n,\\
p_i=F_i(q_i,\ \alpha_i,\ h)\quad(i=1,\ \ldots,\ n)\\
(\alpha_1+\alpha_2+\ldots+\alpha_n=0)
\end{array}\right\} \tag{40}$$

As a special case we get **Liouville's theorem:**
If the kinetic and potential energies of a system have the form

$$T=\frac{1}{2}\left(\sum_{i=1}^{n} A_i\right)\sum_{i=1}^{n} B_i\dot{q}_i^2,$$

$$\Pi=\frac{\sum\limits_{i=1}^{n}\Pi_i}{\sum\limits_{i=1}^{n} A_i} \tag{41}$$

where A_i, B_i and Π_i are functions of one variable q_i $(i=1, \ldots, n)$, then the final equations of motion of the system may be obtained by quadratures.

Indeed, for a Liouville system

$$H=\frac{\sum\limits_{i=1}^{n}\left(\frac{p_i^2}{B_i}+2\Pi_i\right)}{2\sum\limits_{i=1}^{n} A_i} \tag{42}$$

but this is a special case of formula (32).

28. Applying Canonical Transformations to Perturbation Theory

Let the motions of a system with a given function H be known, that is, we know the solutions

$$\left.\begin{array}{l} q_i=\varphi_i(t,\ q_k^0,\ p_k^0),\ p_i=\psi_i(t,\ q_k^0,\ p_k^0) \\ {[q_i^0=(q_i)_{t=0},\ p_i^0=(p_i)_{t=0};\ i=1, \ldots, n]} \end{array}\right\} \tag{1}$$

of the system of differential equations

$$\frac{dq_i}{dt}=\frac{\partial H}{\partial p_i},\quad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q_i}\quad (i=1, \ldots, n) \tag{2}$$

and let it be required to determine the motion of a "perturbed" system with Hamiltonian function $H+H_1$; i.e. let it be required to determine the solutions of the system of differential equations

$$\frac{dq_i}{dt}=\frac{\partial(H+H_1)}{\partial p_i},\quad \frac{dp_i}{dt}=\frac{\partial(H+H_1)}{\partial q_i}\quad (i=1, \ldots, n) \tag{3}$$

If in formulas (1) we regard q_k^0 and p_k^0 $(k=1, \ldots, n)$ as new variables, then, as has been elucidated on page 139, formulas (1) determine a free univalent canonical transformation. This transformation turns the Hamiltonian system (2) into a Hamiltonian system with function $\widetilde{H} = 0$ [see formula (13) on page 138]

$$\frac{dq_k^0}{dt} = 0, \quad \frac{dp_k^0}{dt} = 0 \qquad (k = 1, \ldots, n) \tag{4}$$

and the Hamiltonian system (3) into the Hamiltonian system with function $\widetilde{H}$, which will be equal to H_1:*

$$\frac{dq_k^0}{dt} = \frac{\partial H_1}{\partial p_k^0}, \quad \frac{dp_k^0}{dt} = -\frac{\partial H_1}{\partial q_k^0} \tag{5}$$
$$(k = 1, \ldots, n)$$

Thus, the new variables q_k^0 and p_k^0 possess the following remarkable property; for unperturbed motion, they preserve constant values, which are equal to the initial values; for perturbed motion, they are functions of the time and of the initial values

$$q_k^{[0]}(t, q_j^0, p_j^0), \quad p_k^{[0]}(t, q_j^0, p_j^0) \tag{6}$$
$$(k = 1, \ldots, n)$$

defined as a general solution of the Hamiltonian system (5), in which the "perturbation energy" H_1 is the Hamiltonian function. The final equations for perturbed motion in the initial coordinates q_i, p_i $(i = 1, \ldots, n)$ are obtained by substituting the functions (6) into the formulas for unperturbed motion (1) in place of the constants q_k^0 and p_k^0.

We have succeeded, through the use of the theory of canonical transformations, to replace integration of the Hamiltonian system (3) by integration of the Hamiltonian systems (2) and (5). We obtain the general solution of the system (3) from the general solutions (1) and (6) of these systems by means of superposition:

$$\left.\begin{aligned} q_i &= \varphi_i\,[t,\ q_k^{[0]}(t,\ q_j^0,\ p_j^0),\ p_k^{[0]}(t,\ q_j^0,\ p_j^0)] \\ p_i &= \psi_i\,[t,\ q_k^{[0]}(t,\ q_j^0,\ p_j^0),\ p_k^{[0]}(t,\ q_j^0,\ p_j^0)] \\ & (i = 1, \ldots, n) \end{aligned}\right\} \tag{7}$$

Actually, we have shown that "perturbation in the energy" of a system is equivalent to "perturbation of the initial conditions". This is illustrated in Fig. 40.

In the hyperplane $t = 0$, of an extended phase space, take a fixed point M_0 and draw through it an unperturbed straight-line path. i.e. a straight-line path (1) for the system (2). In Fig. 40, this is depicted as the heavy line M_0N_0. The light line M_0M_0' depicts the displacement, in the hyperplane $t=0$, of the initial point given by the functions (6). From point M_{0t} of this curve draw an unperturbed straight-line path $M_{0t}N_{0t}$ (it is shown dashed in Fig. 40). On this path take a point P with a given value of the time coordinate t. This will then be the position of the system in perturbed motion at time t. In the case of unperturbed motion, the system occupied the position Q at time t. Thus, the perturbation appeared in the "shift" QP. The straight-line path in perturbed motion is shown by the heavy line M_0P. Thus, perturbed motion may be regarded as "compound" motion in phase space: a point is in motion along an unperturbed straight-line path but the path itself is displaced (in the general case it is deformed) due to the "perturbation" of the initial conditions.

---

* This follows from the relation $\widetilde{H} - (H + H_1) = 0 - H$. The left and right sides of this equation are equal to $\frac{\partial S}{\partial t}$, where S is the generating function of the free univalent canonical transformation under consideration (see p. 133).

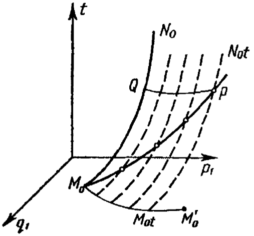

Fig. 40

29. The Structure of an Arbitrary Canonical Transformation

In this and the subsequent sections of this chapter we shall give certain additional information about canonical transformations.

For an arbitrary canonical transformation it is possible to establish formulas that define it with the aid of a generating function and valence c, as was done in Sec. 25 for a free canonical transformation.

For example, of the $4n$ quantities

$$q_i,\ p_i,\ \tilde{q}_i,\ \tilde{p}_i \qquad (i=1,\ \ldots,\ n) \tag{1}$$

connected by the canonical transformation

$$\tilde{q}_i=\varphi_i(t,\ q_k,\ p_k),\quad \tilde{p}_i=\psi_i(t,\ q_k,\ p_k)$$
$$\left(i=1,\ \ldots,\ n;\ \frac{\partial(\tilde{q}_1,\ \ldots,\ \tilde{p}_n)}{\partial(q_1,\ \ldots,\ p_n)}\neq 0\right) \tag{2}$$

it is possible to take the following quantities as $2n$ independent ones:

$$q_1, \ldots, q_l, \; p_{l+1}, \ldots, p_n, \; \tilde{q}_1, \ldots, \; \tilde{q}_m, \; \tilde{p}_{m+1}, \ldots, \tilde{p}_n$$
$$(0 \leqslant l, \; m \leqslant n) \tag{3}$$

Then with the aid of the identities

$$\sum_{g=l+1}^{n} p_g \delta q_g = \delta \Big(\sum_{g=l+1}^{n} q_g p_g \Big) - \sum_{g=l+1}^{n} q_g \delta p_g,$$
$$\sum_{h=m+1}^{n} \tilde{p}_h \delta \tilde{q}_h = \delta \Big(\sum_{h=m+1}^{n} \tilde{q}_h \tilde{p}_h \Big) - \sum_{h=m+1}^{n} \tilde{g}_h \delta \tilde{p}_h$$

we can write the basic defining equation (9) on page 130 in the form

$$\sum_{j=1}^{m} \tilde{p}_j \delta \tilde{q}_j - \sum_{h=m+1}^{n} \tilde{q}_h \delta \tilde{p}_h - \tilde{H} \delta t =$$
$$= c \Big(\sum_{i=1}^{l} p_i \delta q_i - \sum_{g=l+1}^{n} q_g \delta p_g - H \delta t \Big) - \delta U \tag{4}$$

where

$$U = F + \sum_{h=m+1}^{n} \tilde{q}_h \tilde{p}_h - c \sum_{g=l+1}^{n} q_g p_g \tag{5}$$

Since all the $4n$ quantities of (1) are expressed in terms of the $2n$ quantities of (3), we may take it that U is a function of the quantities of (3). Then from the identity (4) we at once find

$$\frac{\partial U}{\partial q_i} = c p_i, \quad \frac{\partial U}{\partial p_g} = -c q_g, \quad \frac{\partial U}{\partial \tilde{q}_j} = -\tilde{p}_j,$$
$$\frac{\partial U}{\partial \tilde{p}_h} = \tilde{q}_h \tag{6}$$
$$(i = 1, \ldots, l; \; g = l + 1, \ldots, n; \; j = 1. \ldots, m;$$
$$h = m + 1, \ldots, n)$$

The formulas (6) are equivalent to the formulas (2) and determine the canonical transformation under consideration with the aid of the valence c and the generating function U of the independent quantities (3).

Below we shall prove a mathematical lemma, according to which it is always possible to choose $2n$ independent quantities of the $4n$ quantities (1) connected by the transformation (2) in such a way that there will not be a single pair of conjugate q_i, p_i or $\tilde{q}_i$, $\tilde{p}_i$ among the quantities chosen. Then if we appropriately renumber the coordinates q_i, $\tilde{q}_k$ $(i, k = 1, \ldots, n)$ and do the same with the momenta p_i, $\tilde{p}_k$ $(i, k = 1, \ldots, n)$, the chosen $2n$ independent quantities may be given in the form (3). Therefore, for an arbitrary canonical transformation there are formulas that can differ from the formulas (6) solely in the numbering of the quantities of (1).*

As was done in Sec. 25 for a free canonical transformation, it may also be shown that for an arbitrary canonical transformation a determinant of order n composed of mixed second-order derivatives of the generating function U is different from zero.** For this reason, the first n equations (6) may be solved for the quantities $\tilde{q}_j$, $\tilde{p}_h$ $(j = 1, \ldots, m;\ h = m + 1, \ldots, n)$. After substituting the expression obtained into the last n equations of (6) we can represent the equations of the canonical transformation (6) in the form of (2).

Let us state and prove the lemma which we used to obtain the structural formulas (6) for an arbitrary canonical transformation.

Lemma. If there are given $2n$ independent functions $\tilde{q}_1, \ldots, \tilde{q}_n$, $\tilde{p}_1, \ldots, \tilde{p}_n$ of $2n$ independent quantities $q_1, \ldots, q_n, p_1, \ldots, p_n$, then out of $4n$ quantities q_i, p_i, $\tilde{q}_i$, $\tilde{p}_i$ $(i = 1, \ldots, n)$ it is always possible to choose $2n$ independent ones in such a manner that there will not be a single pair of conjugate ones (q_k, p_k) or $(\tilde{q}_k, \tilde{p}_k)$ among them.

Proof. Assume the converse, that is, that any $2n$ of the quantities under consideration, among which there is not a single pair of conjugates, are always dependent. Then choose $n + d$ independent ones of the given $4n$ quantities so that the first n do not have the sign $\sim$, and the last d have this sign; choose these $n + d$ quantities so that there are no conjugate ones among them and so that the number d has the greatest of all possible values. By assumption $d < n$.

Since, without altering the conditions or the statement of the lemma it is possible to interchange the roles of the two conjugate quantities q_i and p_i or $\tilde{q}_i$ and $\tilde{p}_i$ and also to perform an arbitrary permutation of the indices $1, \ldots, n$ both of the quantities q_i, p_i

---

* This proposition is given by Carathéodory in his book *Variationsrechnung und partielle Differentialgleichungen erstes Ordnung*, 2 Aufl., Bd. I, 1956, Sec. 96). However, the lemma on which our proof is based (see below) was established by Carathéodory for the special case when the transition from the variables q_i, p_i to the variables $\tilde{q}_i$, $\tilde{p}_i$ $(i = 1, \ldots, n)$ is a canonical transformation.

** This condition may be written as follows: $\det\left(\frac{\partial^2 U}{\partial r_i\, \partial \tilde{r}_k}\right)_{i,k=1}^{n} \neq 0$ if by $r_1, \ldots, r_n, \tilde{r}_1, \ldots, \tilde{r}_n$ we denote, respectively, the quantities (3) in the order given above [see (3)].

and $\tilde{q}_i, \tilde{p}_i$, it may be assumed that the chosen ones are the following $n+d$ quantities:

$$q_1, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d \quad (d < n) \tag{7}$$

We shall call the quantities of (7) the maximum basis. It is obvious that

$$\begin{aligned} \tilde{q}_j &= f(q_1, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d), \\ & \qquad\qquad\qquad\qquad (j = d+1, \ldots, n) \\ \tilde{p}_j &= f(q_1, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d) \end{aligned} \tag{8}$$

where f is the functional symbol.*

Without loss of generality, it may be taken that of the quantities

$$p_1, \ldots, p_n, \tilde{p}_1, \ldots, \tilde{p}_d$$

that do not enter into the formulas (8), the quantities

$$p_1, \ldots, p_a, \tilde{p}_1, \ldots, \tilde{p}_b \quad (a \leqslant n,\ b \leqslant d) \tag{9}$$

are functions of the basal quantities (7), and each of the quantities

$$p_{a+1}, \ldots, p_n, \tilde{p}_{b+1}, \ldots, \tilde{p}_d \tag{10}$$

is independent relative to the basis. Thus,

$$\left.\begin{aligned} p_i &= f(q_1, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d) \quad (i = 1, \ldots, a), \\ \tilde{p}_k &= f(q_1, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d) \quad (k = 1, \ldots, b) \end{aligned}\right\} \tag{11}$$

We shall now show that the quantities $q_{a+1}, \ldots, q_n, \tilde{q}_{b+1}, \ldots, \tilde{q}_d$ which are conjugate to the "independent" quantities of (10), actually do not appear on the right-hand sides of (8). Indeed, let $q_\lambda\ (\lambda > a)$ enter into, say, the expression for a certain $\tilde{q}_j\ (j > d)$:

$$\tilde{q}_j = f(\ldots, q_\lambda, \ldots)$$

Then the system of quantities $q_1, \ldots, q_{\lambda-1}, q_{\lambda+1}, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d, \tilde{q}_j$ is equivalent to the basis (7):

$$\begin{aligned} &(q_1, \ldots, q_{\lambda-1}, q_{\lambda+1}, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d, \tilde{q}_j) \backsim \\ &\backsim (q_1, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d) \end{aligned} \tag{12}$$

---

[*] In different formulas, the same letter f is used to denote different functional relations.

that is, every one of the quantities that appear in one of the two systems is expressed in terms of the quantities of the other and vice versa.* Therefore, the $n+d+1$ quantities

$$q_1, \ldots, q_{\lambda-1}, p_\lambda, q_{\lambda+1}, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d, \tilde{q}_j$$

are independent, which contradicts the "maximality" of the basis (7).Thus, the formulas (8) may be written as

$$\begin{gathered} \tilde{q}_j = f(q_1, \ldots, q_a, \tilde{q}_1, \ldots, \tilde{q}_b) \\ \tilde{p}_j = f(q_1, \ldots, q_a, \tilde{q}_1, \ldots, \tilde{q}_b) \\ (j = d+1, \ldots, n) \end{gathered} \tag{13}$$

Denote by $q_1, \ldots, q_{a_1}, \tilde{q}_1, \ldots, \tilde{q}_{b_1}$ $(a_1 \leqslant a,\ b_1 \leqslant b)$ all the quantities that actually appear at least in one of the right-hand sides of the formulas (13); then

$$\begin{gathered} \tilde{q}_j = f(q_1, \ldots, q_{a_1}, \tilde{q}_1, \ldots, \tilde{q}_{b_1}) \\ \tilde{p}_j = f(q_1, \ldots, q_{a_1}, \tilde{q}_1, \ldots, \tilde{q}_{b_1}) \\ (j = d+1, \ldots, n) \end{gathered} \tag{14}$$

Now we shall show that the quantities q_λ, $\tilde{q}_\mu$ $(\lambda > a,\ \mu > b)$ actually do not appear in those formulas (11) where $i \leqslant a_1$, and $k \leqslant b_1$. Assume the contrary. For instance, let q_λ $(\lambda > a)$ actually appear in the expression (11) for p_{i_1}, where $i_1 \leqslant a_1$,

$$p_{i_1} = f(\ldots, q_\lambda, \ldots) \qquad (i_1 \leqslant a_1,\ \lambda > a)$$

But the quantity q_{i_1} actually appears in one of the expressions (14); for example, let it appear in the expression for $\tilde{q}_j$ $(j > d)$:

$$\tilde{q}_j = f(\ldots, q_{i_1}, \ldots) \qquad (i_1 \leqslant a_1,\ j > d)$$

Then

$$\begin{gathered} (q_1, \ldots, q_{i_1-1}, p_{i_1}, q_{i_1+1}, \ldots, q_{\lambda-1}, q_{\lambda+1}, \ldots \\ \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d, \tilde{q}_j) \backsim (q_1, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d) \end{gathered} \tag{15}$$

and hence the $n+d+1$ quantities

$$\begin{gathered} q_1, \ldots, q_{i_1-1}, p_{i_1}, q_{i_1+1}, \ldots, q_{\lambda-1}, p_\lambda, \\ q_{\lambda+1}, \ldots, q_n, \tilde{q}_1, \ldots, \tilde{q}_d, \tilde{q}_j \end{gathered}$$

---

* If q_λ actually does appear in some functional relation, this relation may be solved for q_λ.

are independent, which fact contradicts the "maximality" of the basis (7).

Thus,

$$\begin{gathered}
p_{i_1}=f\,(q_1,\,\ldots,\,q_a,\;\tilde{q}_1,\,\ldots,\,\tilde{q}_b)\\
\tilde{p}_{k_1}=f\,(q_1,\,\ldots,\,q_a,\;\tilde{q}_1,\,\ldots,\,\tilde{q}_b)\\
(i_1=1,\,\ldots,\,a_1,\;\;k_1=1,\,\ldots,\,b_1)
\end{gathered}\tag{16}$$

Denote as $q_{a_1+1},\,\ldots,\,q_{a_2},\,\tilde{q}_{b_1+1},\,\ldots,\,\tilde{q}_{b_2}$ those of the quantities q_i, $\tilde{q}_k$ $(i>a_1,\;k>b_1)$ which actually occur in the right members of formulas (16). Then

$$\left.\begin{aligned}
p_{i_1}&=f\,(q_1,\,\ldots,\,q_{a_2},\,\tilde{q}_1,\,\ldots,\,\tilde{q}_{b_2})\quad(i_1=1,\,\ldots,\,a_1;\;\;a_1\leqslant a_2\leqslant a)\\
\tilde{p}_{k_1}&=f\,(q_1,\,\ldots,\,q_{a_2},\,\tilde{q}_1,\,\ldots,\,\tilde{q}_{b_2})\quad(k_1=1,\,\ldots,\,b_1;\;\;b_1\leqslant b_2\leqslant b)
\end{aligned}\right\}\tag{17}$$

We shall now show that the quantities q_λ, $\tilde{q}_\mu$ $(\lambda>a,\;\mu>b)$ actually do not occur in the expressions (11) for p_i, $\tilde{p}_k$ where $i\leqslant a_2$, $k\leqslant b_2$. Indeed, let $q_\lambda\,(\lambda>a)$ for example actually appear in the expression for $p_{i_2}\,(a_1<i_2\leqslant a_2)$:

$$p_{i_2}=f\,(\ldots,\,q_\lambda,\,\ldots)\qquad(a_1<i_2\leqslant a_2,\;\lambda>a)$$

But the quantity q_{i_2} actually occurs in one of the expressions (17), for instance in the expression for p_{i_1}. Then the quantity q_{i_1} actually occurs in one of the expressions (14), for instance in the expression for $\tilde{q}_j\,(j>d)$:

$$\begin{aligned}
p_{i_1}&=f\,(\ldots,\,q_{i_2},\,\ldots)\qquad(i_1\leqslant a_1<i_2\leqslant a_2)\\
\tilde{q}_j&=f\,(\ldots,\,q_{i_1},\,\ldots)\qquad(i_1\leqslant a_1,\;j>d)
\end{aligned}$$

In this case an equivalence is established between the system of quantities

$$\begin{gathered}
q_1,\,\ldots,\,q_{i_1-1},\;p_{i_1},\;q_{i_1+1},\,\ldots,\,q_{i_2-1},\;p_{i_2},\;q_{i_2+1},\,\ldots\\
\ldots,\,q_{\lambda-1},\;q_{\lambda+1},\,\ldots,\,q_n,\;\tilde{q}_1,\,\ldots,\,\tilde{q}_d,\;\tilde{q}_j
\end{gathered}\tag{18}$$

and the basis (7). For this reason, the quantity p_λ is independent of the quantities of (18). Adding p_λ to the quantities of (18) we get a basis of $n+d+1$ quantities, which is impossible.

Thus,

$$\begin{gathered}
p_{i_2}=f\,(q_1,\,\ldots,\,q_a,\,\tilde{q}_1,\,\ldots,\,\tilde{q}_b)\\
\tilde{p}_{k_2}=f\,(q_1,\,\ldots,\,q_a,\,\tilde{q}_1,\,\ldots,\,\tilde{q}_b)\\
(i_2=a_1+1,\,\ldots,\,a_2,\;\;k_2=b_1+1,\,\ldots,\,b_2)
\end{gathered}\tag{19}$$

Denote by $q_{a_2+1}, \ldots, q_{a_3}, \tilde{q}_{b_2+1}, \ldots, \tilde{q}_{b_3}$ those of the quantities $q_i, \tilde{q}_k$ $(i > a_2, k > b_2)$ which actually enter into the formulas (19). Then (19) will be written as

$$\left.\begin{aligned} p_{i_2} &= f(q_1, \ldots, q_{a_3}, \tilde{q}_1, \ldots, \tilde{q}_{b_3}) \\ (i_2 &= a_1 + 1, \ldots, a_2;\ a_1 \leqslant a_2 \leqslant a_3 \leqslant a) \\ \tilde{p}_{k_2} &= f(q_1, \ldots, q_{a_3}, \tilde{q}_1, \ldots, \tilde{q}_{b_3}) \\ (k_2 &= b_1 + 1, \ldots, b_2;\ b_1 \leqslant b_2 \leqslant b_3 \leqslant b) \end{aligned}\right\} \tag{20}$$

This process may be continued until the equalities $a_s = a_{s+1}$, $b_s = b_{s+1}$ are attained simultaneously; then

$$\left.\begin{aligned} p_{i_s} &= f(q_1, \ldots, q_{a_s}, \tilde{q}_1, \ldots, \tilde{q}_{b_s}) \\ (i_s &= a_{s-1} + 1, \ldots, a_s;\ a_1 \leqslant \ldots \leqslant a_s \leqslant a) \\ \tilde{p}_{k_s} &= f(q_1, \ldots, q_{a_s}, \tilde{q}_1, \ldots, \tilde{q}_{b_s}) \\ (k_s &= b_{s-1} + 1, \ldots, b_s;\ b_1 \leqslant \ldots \leqslant b_s \leqslant b) \end{aligned}\right\} \tag{21}$$

In place of the formulas (14), (17), (20), (21) we can write

$$\begin{aligned} \tilde{q}_j &= f(q_1, \ldots, q_{a_s}, \tilde{q}_1, \ldots, \tilde{q}_{b_s}) \\ \tilde{p}_j &= f(q_1, \ldots, q_{a_s}, \tilde{q}_1, \ldots, \tilde{q}_{b_s}) \end{aligned} \quad (j = d + 1, \ldots, n) \tag{22}$$

$$p_i = f(q_1, \ldots, q_{a_s}, \tilde{q}_1, \ldots, \tilde{q}_{b_s}) \quad (i = 1, \ldots, a_s) \tag{23}$$

$$\tilde{p}_k = f(q_1, \ldots, q_{a_s}, \tilde{q}_1, \ldots, \tilde{q}_{b_s}) \quad (k = 1, \ldots, b_s) \tag{24}$$

Now let $a_s > b_s$. Then from the formulas (23) we can eliminate $\tilde{q}_1, \ldots, \tilde{q}_{b_s}$ and obtain a relation between q_i and p_i; but this contradicts the hypothesis of the lemma.

If $a_s \leqslant b_s$, then $a_s < b_s + 1$. Then from the formulas (22) and (24) we can eliminate all the q_i and obtain a relation between $\tilde{q}_k$ and $\tilde{p}_k$; this again contradicts the hypothesis.

Thus, the assumption that there exists a maximal basis (7) in which $d < n$ has led us to a contradiction. The lemma is proved.

30. Testing the Canonical Character of a Transformation. The Lagrange Brackets

Let us establish certain tests for the canonical character of a transformation, i.e. the necessary and sufficient conditions that must be satisfied by $2n$ independent functions [with respect to q_k, p_k $(k = 1, \ldots, n)$]

$$\tilde{q}_i = \varphi_i(t, q_k, p_k), \quad \tilde{p}_i = \psi_i(t, q_k, p_k) \quad (i = 1, \ldots, n) \tag{1}$$

so that the transformation defined by these functions should be canonical.

Let the transformation (1) be canonical. We write out the defining identity for it:

$$\sum_{i=1} \tilde{p}_i \delta \tilde{q}_i - \tilde{H} \delta t = c \left(\sum_{i=1}^{n} p_i \delta q_i - H \delta t \right) - \delta F(t, q_i, p_i) \tag{2}$$

Take an arbitrary fixed value $t = \bar{t}$. Then from the identity (2) we get

$$\sum_{i=1}^{n} \tilde{p}_i \delta \tilde{q}_i = c \sum_{i=1}^{n} p_i \delta q_i - \delta F(\bar{t}, q_i, p_i) \tag{3}$$

But (3) is a defining identity for a transformation that does not contain the time explicitly,

$$\tilde{q}_i = \varphi_i(\bar{t}, q_k, p_k), \ \tilde{p}_i = \psi_i(\bar{t}, q_k, p_k) \quad (i = 1, \ldots, n) \tag{4}$$

Hence, formulas (4) define a canonical transformation with valence c which is independent of the chosen value of $t = \bar{t}$.

On the contrary, let it now be given that all transformations obtained from the transformation (1) by replacing the variable t by various fixed values of $\bar{t}$ are canonical and with one and the same valence c. Then, defining the function $\tilde{H}$ by the equation

$$\tilde{H} = cH + \frac{\partial F}{\partial t} + \sum_{i=1}^{n} \tilde{p}_i \frac{\partial \tilde{q}_i}{\partial t} \tag{5}$$

we get equation (2) from (3) and (5), i.e. we find that the transformation (1) that depends on the time t is canonical.

Thus, for the time-dependent transformation (1) to be canonical it is necessary and sufficient that all the time-independent transformations obtained from the transformation (1) by replacing t with an arbitrary value of $\bar{t}$ be canonical and with one and the same valence c.

For this reason, when establishing tests for canonical character, we can confine ourselves to canonical transformations that do not contain the time variable t explicitly:

$$\tilde{q}_i = \varphi_i(q_k, p_k), \quad \tilde{p}_i = \psi_i(q_k, p_k)$$
$$\left(i = 1, \ldots, n; \ \frac{\partial(\tilde{q}_1, \ldots, \tilde{p}_n)}{\partial(q_1, \ldots, p_n)} \neq 0 \right) \tag{6}$$

For the canonical transformation (6), the defining identity (2) is written as

$$\sum_{k=1}^{n} \tilde{p}_k \delta \tilde{q}_k = c \sum_{k=1}^{n} p_k \delta q_k - \delta K(q_k, p_k) \tag{7}$$

Here, express $\delta\tilde{q}_k$ in terms of δq_i and δp_i using formulas (1). Then (7) will take the form

$$\sum_{i=1}^{n} (\Phi_i \delta q_i + \Psi_i \delta p_i) = -\delta K(q_k, p_k) \tag{8}$$

where

$$\Phi_i = \sum_{k=1}^{n} \tilde{p}_k \frac{\partial \tilde{q}_k}{\partial q_i} - c p_i, \quad \Psi_i = \sum_{k=1}^{n} \tilde{p}_k \frac{\partial \tilde{q}_k}{\partial p_i} \quad (i = 1, \ldots, n) \tag{8'}$$

It remains to write out the conditions that the left-hand side of (8) should be differential and we will have a test for canonical character in the form of the equalities

$$\begin{gathered} \frac{\partial \Phi_i}{\partial q_k} = \frac{\partial \Phi_k}{\partial q_i}, \quad \frac{\partial \Psi_i}{\partial p_k} = \frac{\partial \Psi_k}{\partial p_i}, \\ \frac{\partial \Phi_i}{\partial p_k} = \frac{\partial \Psi_k}{\partial q_i} \quad (i, k = 1, \ldots, n) \end{gathered} \tag{9}$$

Substituting the expressions (8′) into these equalities we find (after elementary transformations)

$$\left.\begin{aligned} \sum_{j=1}^{n} \left(\frac{\partial \tilde{q}_j}{\partial q_i} \frac{\partial \tilde{p}_j}{\partial q_k} - \frac{\partial \tilde{q}_j}{\partial q_k} \frac{\partial \tilde{p}_j}{\partial q_i} \right) &= 0 \\ \sum_{j=1}^{n} \left(\frac{\partial \tilde{q}_j}{\partial p_i} \frac{\partial \tilde{p}_j}{\partial p_k} - \frac{\partial \tilde{q}_j}{\partial p_k} \frac{\partial \tilde{p}_j}{\partial p_i} \right) &= 0 \\ \sum_{j=1}^{n} \left(\frac{\partial \tilde{q}_j}{\partial q_i} \frac{\partial \tilde{p}_j}{\partial p_k} - \frac{\partial \tilde{q}_j}{\partial p_k} \frac{\partial \tilde{p}_j}{\partial q_i} \right) &= c\delta_{ik} \quad (i, k = 1, \ldots, n) \end{aligned}\right\} \tag{10}$$

where δ_{ik} is Kronecker's symbol: $\delta_{ik} = 0$ for $i \neq k$; $\delta_{ik} = 1$ for $i = k$ $(i, k = 1, \ldots, n)$.

The conditions (10) may be written in compact form if we introduce what are called *Lagrange brackets*, which, for the given $2n$ functions φ_i, ψ_i $(i = 1, \ldots, n)$ of the two variables q and p are defined as follows:*

$$[qp] = \sum_{j=1}^{n} \left(\frac{\partial \varphi_j}{\partial q} \frac{\partial \psi_j}{\partial p} - \frac{\partial \varphi_j}{\partial p} \frac{\partial \psi_j}{\partial q} \right) = \sum_{j=1}^{n} \frac{\partial (\varphi_j, \psi_j)}{\partial (q, p)} \tag{11}$$

---

* The reader can compare the Lagrange brackets with the Poisson brackets introduced in Sec. 15. There, two functions φ, ψ were given of $2n$ variables q_i, p_i and the Poisson brackets equaled the sum of the Jacobians $\frac{\partial(\varphi, \psi)}{\partial(q_i, p_i)}$. Here, however, $2n$ functions of two arguments are given and the Lagrange brackets are equal to the sum of the Jacobians (11).

Using this notation and taking for the φ_i, ψ_i the functions $\tilde{q}_i$, $\tilde{p}_i$ $(i=1, \ldots, n)$ defined by the formulas (6), we can write the conditions (10) as

$$[q_i q_k]=0, \quad [p_i p_k]=0, \quad [q_i p_k]=c\delta_{ik} \quad (i, k=1, \ldots, n) \tag{12}$$

where c is the valence of the canonical transformation.

The equalities (12) express the necessary and sufficient conditions for the transformation (6) to be canonical. In the case of a time-dependent transformation, the conditions (12) are preserved, provided that they hold for any value of t.

31. The Simplicial Nature of the Jacobian Matrix of a Canonical Transformation

Consider the Jacobian matrix of the canonical transformation

$$\boldsymbol{M}=\left\| \begin{matrix} \frac{\partial \tilde{q}_1}{\partial q_1} & \ldots & \frac{\partial \tilde{q}_1}{\partial q_n} & \frac{\partial \tilde{q}_1}{\partial p_1} & \ldots & \frac{\partial \tilde{q}_1}{\partial p_n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{\partial \tilde{q}_n}{\partial q_1} & \ldots & \frac{\partial \tilde{q}_n}{\partial q_n} & \frac{\partial \tilde{q}_n}{\partial p_1} & \ldots & \frac{\partial \tilde{q}_n}{\partial p_n} \\ \frac{\partial \tilde{p}_1}{\partial q_1} & \ldots & \frac{\partial \tilde{p}_1}{\partial q_n} & \frac{\partial \tilde{p}_1}{\partial p_1} & \ldots & \frac{\partial \tilde{p}_1}{\partial p_n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{\partial \tilde{p}_n}{\partial q_1} & \ldots & \frac{\partial \tilde{p}_n}{\partial q_n} & \frac{\partial \tilde{p}_n}{\partial p_1} & \ldots & \frac{\partial \tilde{p}_n}{\partial p_n} \end{matrix} \right\| = \begin{pmatrix} \frac{\partial \tilde{\boldsymbol{q}}}{\partial \boldsymbol{q}} & \frac{\partial \tilde{\boldsymbol{q}}}{\partial \boldsymbol{p}} \\ \frac{\partial \tilde{\boldsymbol{p}}}{\partial \boldsymbol{q}} & \frac{\partial \tilde{\boldsymbol{p}}}{\partial \boldsymbol{p}} \end{pmatrix} \tag{1}$$

Here $\frac{\partial \tilde{\boldsymbol{q}}}{\partial \boldsymbol{q}}$ is a Jacobian matrix of order n $\left\| \frac{\partial \tilde{q}_i}{\partial q_k} \right\|$. Similarly, we define the n-order Jacobian matrices $\frac{\partial \tilde{\boldsymbol{q}}}{\partial \boldsymbol{p}}$, $\frac{\partial \tilde{\boldsymbol{p}}}{\partial \boldsymbol{q}}$ and $\frac{\partial \tilde{\boldsymbol{p}}}{\partial \boldsymbol{p}}$.

We introduce a special matrix of order $2n$:

$$\boldsymbol{J}=\left\| \begin{matrix} 0 & \ldots & 0 & -1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & 0 & 0 & \ldots & -1 \\ 1 & \ldots & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & 1 & 0 & \ldots & 0 \end{matrix} \right\| = \begin{pmatrix} 0 & -\boldsymbol{E} \\ \boldsymbol{E} & 0 \end{pmatrix} \tag{2}$$

where $\boldsymbol{E}$ is a unit matrix of order n. Consider the matrix $\boldsymbol{M}$ and also its transpose $\boldsymbol{M}'$; form the product $\boldsymbol{M}'\boldsymbol{J}\boldsymbol{M}$ and prove that by virtue of the relation (12) of the preceding section this product is identically equal to cJ:

$$M'JM = cJ \tag{3}$$

where c is the valence of the canonical transformation.

Indeed,*

$$M'J = \begin{pmatrix} \left(\frac{\partial \tilde{q}}{\partial q}\right)' & \left(\frac{\partial \tilde{p}}{\partial q}\right)' \\ \left(\frac{\partial \tilde{q}}{\partial p}\right)' & \left(\frac{\partial \tilde{p}}{\partial p}\right)' \end{pmatrix} \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix} = \begin{pmatrix} \left(\frac{\partial \tilde{p}}{\partial q}\right)' & -\left(\frac{\partial \tilde{q}}{\partial q}\right)' \\ \left(\frac{\partial \tilde{p}}{\partial p}\right)' & -\left(\frac{\partial \tilde{q}}{\partial p}\right)' \end{pmatrix}$$

$$M'JM = \begin{pmatrix} \left(\frac{\partial \tilde{p}}{\partial q}\right)' \frac{\partial \tilde{q}}{\partial q} - \left(\frac{\partial \tilde{q}}{\partial q}\right)' \frac{\partial \tilde{p}}{\partial q} & \left(\frac{\partial \tilde{p}}{\partial q}\right)' \frac{\partial \tilde{q}}{\partial p} - \left(\frac{\partial \tilde{q}}{\partial q}\right)' \frac{\partial \tilde{p}}{\partial p} \\ \left(\frac{\partial \tilde{p}}{\partial p}\right)' \frac{\partial \tilde{q}}{\partial q} - \left(\frac{\partial \tilde{q}}{\partial p}\right)' \frac{\partial \tilde{p}}{\partial q} & \left(\frac{\partial \tilde{p}}{\partial p}\right)' \frac{\partial \tilde{q}}{\partial p} - \left(\frac{\partial \tilde{q}}{\partial p}\right)' \frac{\partial \tilde{p}}{\partial p} \end{pmatrix}$$

But

$$\left(\frac{\partial \tilde{p}}{\partial q}\right)' \frac{\partial \tilde{q}}{\partial q} - \left(\frac{\partial \tilde{q}}{\partial q}\right)' \frac{\partial \tilde{p}}{\partial q} = \left\| \sum_{j=1}^{n} \frac{\partial \tilde{p}_j}{\partial q_i} \frac{\partial \tilde{q}_j}{\partial q_k} - \frac{\partial \tilde{q}_j}{\partial q_i} \frac{\partial \tilde{p}_j}{\partial q_k} \right\|_{i,\,k=1}^{n} = \| [q_i q_k] \| = 0$$

Performing similar computations for the remaining three blocks, we get

$$M'JM = \begin{pmatrix} 0 & -cE \\ cE & 0 \end{pmatrix} = cJ$$

which is what we sought to prove.

For a univalent canonical transformation, equation (3) is written as

$$M'JM = J \tag{4}$$

The matrices M for which the equation (4) holds are called simplicial. Since $\det J = 1$, and the determinant of a product of matrices is equal to the product of the determinants of the matrices being multiplied, we find from (4) that

$$\det M = \pm 1$$

Thus, simplicial matrices are nonsingular.**

---

* When the elements of the matrix are matrix blocks, multiplication is performed according to the same rules as when the elements of the matrix are numbers, i.e. the rows of the first matrix factor are multiplied by the columns of the second matrix factor.

** As may readily be verified, the product of two simplicial matrices, the inverse matrix for any simplicial matrix and the unit matrix are again simplicial matrices. Therefore, simplicial matrices form a group, a *simplicial group*. (Continued on p. 163.)

We shall call the matrix M that satisfies the relation (3) a *generalized-simplicial matrix* (with valence c).*

Since the relations (12) of the preceding section reduced to the condition of the generalized-simplicial nature of the Jacobian matrix M, it follows that the test for the canonical character of a transformation may be stated thus:

For a certain transformation $\tilde{q}_i = \tilde{q}_i(t, q_k, p_k)$, $\tilde{p}_i = \tilde{p}_i(t, q_k, p_k)$ $(i = 1, \ldots, n)$ to be canonical, it is necessary and sufficient that the Jacobian matrix M corresponding to this transformation should be generalized-simplicial with constant valence c. (In the case of a univalent transformation, the matrix M is an ordinary simplicial matrix.) Then, the condition of the simplicial nature of (3) must hold identically relative to all the variables t, q_i, p_i $(i = 1, \ldots, n)$.

In Sec. 26 it was established that the motion of any Hamiltonian system may be regarded as a free univalent canonical transformation. Consequently, its Jacobian matrix is simplicial and has the determinant I (see pages 125-126) equal to ± 1.

Since at the initial time $I(0) = +1$, so $I = +1$ at all subsequent times, but it is precisely this determinant that serves as the integrand in the expression for the phase volume in $2n$-dimensional space [formula (3), Sec. 23].

This remark may be regarded as a proof of the Liouville theorem that differs from the proof given in Sec. 23.

32. Invariance of the Poisson Brackets in a Canonical Transformation

Let us give the condition of canonicity of a transformation written in the form of equation (3) in the preceding section in a somewhat modified form. Multiply both sides of this equation by $(M')^{-1}$ on the left and by M^{-1} on the right. We get

$$(M')^{-1} J M^{-1} = \frac{1}{c} J \tag{1}$$

Simplicial matrices are characterized by the following properties. In the bilinear form $f = \sum_{i=1}^{n} (x_i y_i' - y_i x_i')$, we subject the $2n$ variables x_i, y_i and the $2n$ variables x_i', y_i' to one and the same linear transformation with the simplicial matrix of coefficients M. Then the form f preserves its aspect in the new variables x_i^*, y_i^* and $x_i^{*\prime}$, $y_i^{*\prime}$: $f = \sum_{i=1}^{n} (x_i^* y_i^{*\prime} - y_i^* x_i^{*\prime})$.

* All generalized-simplicial matrices (for all $c \neq 0$) also form a group. If M is a generalized-simplicial matrix, then det $M = \pm c^n$.

Take the inverse matrices of both sides of the equation, noting that $J^{-1}=-J$:*

$$MJM' = cJ \tag{2}$$

The equality $MJM' = cJ$ is obtained from the equality $M'JM = cJ$ by replacing the Jacobian matrix M by its transpose M'. But this substitution [see formula (1), Sec. 31] reduces to replacing the derivatives $\frac{\partial \tilde{q}_i}{\partial q_k}, \frac{\partial \tilde{q}_i}{\partial p_k}, \frac{\partial \tilde{p}_i}{\partial q_k}, \frac{\partial \tilde{p}_i}{\partial p_k}$, respectively, by the derivatives $\frac{\partial \tilde{q}_k}{\partial q_i}, \frac{\partial \tilde{p}_k}{\partial q_i}, \frac{\partial \tilde{q}_k}{\partial p_i}, \frac{\partial \tilde{p}_k}{\partial p_i}$; in other words, in each derivative the letters and indices above and below** are interchanged. Thus, if the equation $M'JM = cJ$ was equivalent to the system of equalities

$$[q_iq_k]=0,\ [p_ip_k]=0,\ [q_ip_k]=c\delta_{ik} \quad (i, k = 1, \ldots, n) \tag{3}$$

then the equation (2) will be equivalent to the system of equalities

$$[q_iq_k]^*=0,\ [p_ip_k]^*=0,\ [q_ip_k]^*=\delta_{ik} \quad (i, k = 1, \ldots, n) \tag{4}$$

where the asterisk (*) indicates that the aforementioned interchange of derivatives is to be performed within the Lagrange brackets. But then the Lagrange brackets become Poisson brackets. Indeed,

$$[q_iq_k]^* = \left[\sum_{j=1}^{n}\left(\frac{\partial \tilde{q}_j}{\partial q_i}\frac{\partial \tilde{p}_j}{\partial q_k} - \frac{\partial \tilde{p}_j}{\partial q_i}\frac{\partial \tilde{q}_j}{\partial q_k}\right)\right]^* =$$

$$= \sum_{j=1}^{n}\left(\frac{\partial \tilde{q}_i}{\partial q_j}\frac{\partial \tilde{q}_k}{\partial p_j} - \frac{\partial \tilde{q}_i}{\partial p_j}\frac{\partial \tilde{q}_k}{\partial q_j}\right) = (\tilde{q}_i\tilde{q}_k)$$

where $(\tilde{q}_i, \tilde{q}_k)$ are Poisson brackets of the functions $\tilde{q}_i$ and $\tilde{q}_k$ with respect to the independent variables $q_1, p_1, \ldots, q_n, p_n$. Similarly,

$$[p_ip_k]^* = (\tilde{p}_i\tilde{p}_k),\quad [q_ip_k]^* = (\tilde{q}_i\tilde{p}_k)$$

Therefore, the conditions of the canonicity of transformation (4) may be written in the following form with the aid of the Poisson brackets:

$$(\tilde{q}_i\tilde{q}_k)=0,\quad (\tilde{p}_i\tilde{p}_k)=0,\quad (\tilde{q}_i\tilde{p}_k)=c\delta_{ik} \quad (i, k = 1, \ldots, n) \tag{5}$$

---

* Here we use the following rule: the inverse matrix of a product of matrices is equal to the product of inverse matrices in reversed order: $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$. Besides,

$$J^2 = \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}\begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix} = -\begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix} = -JJ^{-1}$$

that is, $J^{-1} = -J$.

** As before, the sign ~ remains on top.

Let us now consider two functions φ and ψ of the quantities q_i, p_i $(i=1, \ldots, n)$ and t. Expressing in these functions the q_i, p_i $(i=1, \ldots, n)$ in terms of $\tilde{q}_k$, $\tilde{p}_k$ $(k=1, \ldots, n)$ with the aid of an inverse canonical transformation, we can regard these same functions as functions of the variables $\tilde{q}_k$, $\tilde{p}_k$ $(k=1, \ldots, n)$. Accordingly, the Poisson brackets of φ, ψ may be evaluated both with respect to the variables q_i, p_i [denoted $(\varphi\psi)$] and with respect to the variables $\tilde{q}_i$, $\tilde{p}_i$ [denoted $(\varphi\psi)^{\sim}$].

Let us prove the validity of the identity

$$(\varphi\psi) = c\,(\varphi\psi)^{\sim} \tag{6}$$

Proof of this identity rests on a familiar expression of the Jacobian of a system of composite functions

$$\frac{\partial(\varphi, \psi)}{\partial(q_j, p_j)} = \sum_{\substack{i,k=1\\ i<k}}^{n} \left[\frac{\partial(\varphi, \psi)}{\partial(\tilde{q}_i, \tilde{q}_k)} \frac{\partial(\tilde{q}_i, \tilde{q}_k)}{\partial(q_j, p_j)} + \frac{\partial(\varphi, \psi)}{\partial(\tilde{p}_i, \tilde{p}_k)} \frac{\partial(\tilde{p}_i, \tilde{p}_k)}{\partial(q_j, p_j)}\right] + \sum_{i,k=1}^{n} \frac{\partial(\varphi, \psi)}{\partial(\tilde{q}_i, \tilde{p}_k)} \frac{\partial(\tilde{q}_i, \tilde{p}_k)}{\partial(q_j, p_j)}$$

$$(j=1, \ldots, n)$$

Summing these identities termwise, we get

$$(\varphi\psi) = \sum_{\substack{i,k=1\\ i<k}}^{n} \left[\frac{\partial(\varphi, \psi)}{\partial(\tilde{q}_i, \tilde{q}_k)} (\widetilde{q_i q_k}) + \frac{\partial(\varphi, \psi)}{\partial(\tilde{p}_i, \tilde{p}_k)} (\widetilde{p_i p_k})\right] + \sum_{i,k=1}^{n} \frac{\partial(\varphi, \psi)}{\partial(\tilde{q}_i, \tilde{p}_k)} (\widetilde{q_i p_k})$$

From this, using equalities (5), we find

$$(\varphi\psi) = c \sum_{j=1}^{n} \frac{\partial(\varphi, \psi)}{\partial(\tilde{q}_j, \tilde{p}_j)} = c\,(\varphi\psi)^{\sim} \tag{7}$$

The converse assertion also holds. If for any two functions φ and ψ, the identity (7) is fulfilled for one and the same constant $c \neq 0$, then the transition from the $2n$ variables q_i, p_i to the $2n$ variables $\tilde{q}_i$, $\tilde{p}_i$ is accomplished by a canonical transformation with valence c.

For a univalent canonical transformation $c=1$, and therefore

$$(\varphi\psi) = (\varphi\psi)^{\sim}$$

In other words, the Poisson brackets are invariant under univalent canonical transformations. This property of univalent canonical transformations singles out these transformations from among all possible transformations of phase space.

CHAPTER 5

Stability of Equilibrium and the Motions of a System

33. Lagrange's Theorem on the Stability of an Equilibrium Position

We begin by defining a stable position of equilibrium.

First recall* that a position of a system is called a position of equilibrium if the system, which at the initial time was in that position with zero velocities, remains in that position.

Let the position of a holonomic system be defined by the independent coordinates $q_1, \ldots, q_n$ (where n is the number of degrees of freedom of the system). As was explained in Sec. 5, in a position of equilibrium (and only in that position) all generalized forces are zero: $Q_i = 0$ $(i = 1, \ldots, n)$.** Without loss of generality, we may take it that the position under study of the system lies at the coordinate origin $q_1 = \ldots = q_n = 0$. Then the coordinates of any other position of the system $q_1, \ldots, q_n$ characterize a deviation of this position from the equilibrium position and for this reason are called *deviations* of the system.

The position of equilibrium $q_1 = \ldots = q_n = 0$ (or the state of equilibrium $q_1 = \ldots = q_n = 0$, $\dot{q}_1 = \ldots = \dot{q}_n = 0$) is called *stable* if for sufficiently small initial deviations q_k^0 and sufficiently small initial velocities $\dot{q}_i^0$ $(i = 1, \ldots, n)$ the system does not go beyond the limits of an arbitrary small (preassigned!) neighbourhood of the position of equilibrium during the entire period of motion and has arbitrarily small velocities $\dot{q}_i$ $(i = 1, \ldots, n)$; that is to say, if for any $\varepsilon > 0$ it is possible to indicate a $\delta = \delta(\varepsilon) > 0$ such that for all $t \geqslant t_0$ the inequalities

$$|q_i(t)| < \varepsilon, \quad |\dot{q}_i(t)| < \varepsilon \quad (i = 1, \ldots, n) \tag{1}$$

hold so long as at the initial time $t = t_0$

$$|q_i^0| < \delta, \quad |\dot{q}_i^0| < \delta \quad (i = 1, \ldots, n) \tag{2}$$

---

* See Sec. 4.

** If the generalized forces Q_i depend not only on the coordinates q_k but also on the generalized velocities $\dot{q}_k$ $(k = 1, \ldots, n)$, then the equalities $Q_i = 0$ must hold when the coordinates q_k appearing in the expression for Q_i are replaced by the coordinates of an equilibrium position, and the generalized velocities are assumed equal to zero.

It is convenient to interpret inequalities (1) and (2) geometrically in a $2n$-dimensional state space $(q_i, \dot{q}_i)$. For the case $n = 1$, Fig. 41 depicts, in the plane $(q, \dot{q})$, two neighbourhoods of the coordinate origin O that correspond to the inequalities (1) and (2). If the origin O is a stable state of equilibrium and for a given $\varepsilon > 0$ some $\delta > 0$ is suitably chosen, then any motion starting at time t_0 within a square with centre at O and side 2δ will occur all the time within just such a square with side 2ε.

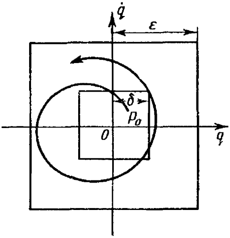

Fig. 41

Example 1. A heavy ball can move along a rim having the shape of a circle and being located in the vertical plane. There are two positions of equilibrium: the lowest and the highest point of the circle. The former is the stable position of equilibrium while the latter is the unstable position of equilibrium.

Example 2. Linear oscillator. The position of equilibrium is stable. Indeed, for a linear oscillator $T = \frac{1}{2} m\dot{q}^2$, $\Pi = \frac{1}{2} cq^2$ $(m > 0,\ c > 0)$, and the differential equation of motion $m\ddot{q} + cq = 0$ has the general solution

$$q = q_0 \cos \omega (t - t_0) + \frac{\dot{q}_0}{\omega} \sin \omega (t - t_0) \quad \left(\omega^2 = \frac{c}{m}\right)$$

Therefore

$$|q(t)| \leqslant |q_0| + \frac{1}{\omega} |\dot{q}_0| < \varepsilon, \quad |\dot{q}(t)| \leqslant \omega |q_0| + |\dot{q}_0| < \varepsilon$$

only if $|q_0| < \delta$, and $|\dot{q}_0| < \delta$ where, for example, $\delta = \min\left(\frac{\varepsilon}{2\omega}, \frac{\omega\varepsilon}{2}\right)$.

Example 3. A particle of mass m can move along the x-axis under the action of two forces: a restoring force proportional to the deviation $-cx$, and a resisting force of the medium, which is proportional to the first power of the velocity $-2f\dot{x}$ $(c > 0,\ f > 0)$.

The point $x = 0$ will be a stable position of equilibrium. We first consider the case when the coefficient of the force of resistance is small: $0 < f < \sqrt{mc}$.

Then the differential equation of motion $m\ddot{x}+2f\dot{x}+cx=0$ has the general solution

$$x=e^{-\frac{f}{m}(t-t_0)}\left[C_1\cos\frac{d}{m}(t-t_0)+C_2\sin\frac{d}{m}(t-t_0)\right]$$

where

$$d=\sqrt{mc-f^2},\quad C_1=x_0,\quad C_2=\frac{1}{d}(fx_0+m\dot{x}_0)$$

From this, for any value of t,

$$|x(t)|\leqslant|C_1|+|C_2|\leqslant\left(1+\frac{f}{d}\right)|x_0|+\frac{m}{d}|\dot{x}_0|<\varepsilon$$

and

$$|\dot{x}(t)|\leqslant\frac{d}{m}\left(1+\frac{f}{d}\right)(|C_1|+|C_2|)\leqslant\frac{d}{m}\left(1+\frac{f}{d}\right)^2|x_0|+\left(1+\frac{f}{d}\right)|\dot{x}_0|<\varepsilon$$

provided that $|x_0|<\delta$, and $|\dot{x}_0|<\delta$, where, for example,

$$\delta=\min\left[\frac{d\varepsilon}{2m},\ \frac{m\varepsilon}{2d\left(1+\frac{f}{d}\right)^2}\right]$$

But if $f>\sqrt{mc}$, then the general solution of the differential equation of motion $m\ddot{x}+2f\dot{x}+cx=0$ has the form

$$x=C_1e^{-\nu_1(t-t_0)}+C_2e^{-\nu_2(t-t_0)}$$

where

$$\nu_{1,2}=\frac{f\pm\sqrt{f^2-mc}}{m}>0,\quad C_1=\frac{\nu_2x_0+\dot{x}_0}{\nu_2-\nu_1},\quad C_2=\frac{\nu_1x_0+\dot{x}_0}{\nu_1-\nu_2}$$

Whence, for any value of t,

$$|x(t)|\leqslant|C_1|+|C_2|\leqslant\frac{(\nu_1+\nu_2)|x_0|+2|\dot{x}_0|}{|\nu_2-\nu_1|}=\frac{f|x_0|+m|\dot{x}_0|}{\sqrt{f^2-mc}}<\varepsilon$$

and

$$|\dot{x}(t)|\leqslant\nu_1|C_1|+\nu_2|C_2|\leqslant\frac{2\nu_1\nu_2|x_0|+(\nu_1+\nu_2)|\dot{x}_0|}{|\nu_2-\nu_1|}=\frac{c|x_0|+f|\dot{x}_0|}{\sqrt{f^2-mc}}<\varepsilon$$

provided that $|x_0|<\delta$, $|\dot{x}_0|<\delta$, where

$$\delta=\min\left(\frac{\sqrt{f^2-mc}}{2f}\varepsilon,\ \frac{\sqrt{f^2-mc}}{2c}\varepsilon,\ \frac{\sqrt{f^2-mc}}{2m}\varepsilon\right)$$

In Examples 2 and 3, the stability of the position of equilibrium was established with the aid of final equations obtained by integrating the differential equations of motion of the system. These final equations of motion gave us the dependence of the deviations of q_i and the generalized velocities $\dot{q}_i$ upon the time t and the initially given q_i^0, $\dot{q}_i^0$ $(i=1, \ldots, n)$. The determination of these final equations of motion and their investigation are extremely involved in more complicated (for example, nonlinear) problems. Of interest therefore are criteria of the stability of an equilibrium position that do not require preliminary integration of the differential equations of motion of the system.

As far back as 1644, Torricelli knew that the position of a system of bodies acted upon by gravity would be stable if the centre of gravity of the system occupied the lowest of possible positions. Lagrange generalized this principle of Torricelli to the case of arbitrary potential forces and established a criterion for the stability of the equilibrium position of a conservative system:

Lagrange's theorem.* *If in some position of a conservative system the potential energy has a strict minimum, then this position is the position of stable equilibrium of the system.*

Proof. Without loss of generality we may assume that all the coordinates $q_1, \ldots, q_n$ and the potential energy $\Pi(q_1, \ldots, q_n)$ are equal to zero in the position under consideration; that is, $q_1 = \ldots = q_n = 0$, and $\Pi(0, \ldots, 0) = 0$.** Since in the given position of the system the function Π has a minimum, the generalized forces in this position are zero:

$$Q_i = -\frac{\partial \Pi}{\partial q_i} = 0 \quad (i = 1, \ldots, n)$$

i.e. the point $q_1 = \ldots = q_n = 0$ is the equilibrium position of the system. Further, from the fact that the value of $\Pi(0, \ldots, 0) = 0$ is a strict minimum it follows that in a certain Δ-neighbourhood of the equilibrium position

$$|q_i| < \Delta \quad (i = 1, \ldots, n) \tag{3}$$

we have a strict inequality

$$\Pi(q_1, \ldots, q_n) > \Pi(0, \ldots, 0) = 0 \tag{4}$$

provided, however, that all coordinates q_i do not vanish at once.

Form the expression for the total energy of the system:

$$E(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n) = T + \Pi =$$

$$= \frac{1}{2} \sum_{i, k=1}^{n} a_{ik}(q_1, \ldots, q_n)\, \dot{q}_i \dot{q}_k + \Pi(q_1, \ldots, q_n) \tag{5}$$

---

* This theorem appeared in Lagrange's *Mécanique Analytique* (1st ed., 1788), but a rigorous proof of the theorem was first given by P. G. Lejeune Dirichlet. For this reason, the theorem is frequently called Dirichlet's theorem.

** The potential energy Π is determined up to an arbitrary additive constant. We choose this constant so that the value of Π should be zero in the equilibrium position.

From the inequality (4) and from the fact that $T > 0$, if at least one of the generalized velocities $\dot{q}_i$ is not equal to zero* it follows that, always,

$$E > 0$$

when (3) holds, provided that all the $2n$ quantities q_i, $\dot{q}_i$ $(i = 1, \ldots, n)$ do not vanish simultaneously; i.e. the *total energy* $E(q_i, \dot{q}_i)$, while vanishing at the origin O of a $2n$-dimensional state space, has a strict minimum (equal to zero) at this point.

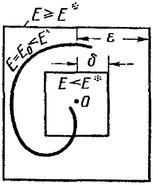

Fig. 42

Now select an arbitrary number ε subject to the sole restriction $0 < \varepsilon < \Delta$; and consider the values of the total energy E on the boundary of the ε-neighbourhood defined by the inequalities

$$|q_i| < \varepsilon, \quad |\dot{q}_i| < \varepsilon \quad (i = 1, \ldots, n) \tag{6}$$

(Fig. 42). Since this boundary is a closed set of points, the continuous function E reaches its minimum E^* on this boundary. Since all values of E are positive on the boundary of the ε-neighbourhood, the minimum E^* also is positive. Thus, on the boundary of the ε-neighbourhood

$$E \geqslant E^* > 0 \tag{7}$$

On the other hand, since the continuous function E vanishes at the origin O, there will always be a δ-neighbourhood of the point O** such that in this neighbourhood

$$E < E^* \tag{8}$$

---

* This is true if the Δ-neighbourhood of the coordinate origin O in the coordinate space does not contain any singular points (see footnote on page 47). We assume that the point O, at which the function Π has a minimum, is not singular. But then there are no singular points in some Δ_0-neighbourhood of point O. We choose $\Delta < \Delta_0$.

** Since the inequality (8) holds in the δ-neighbourhood, $\delta \leqslant \varepsilon$.

Therefore if the initial coordinates and the initial velocities satisfy the inequalities (2), the initial energy $E_0 < E^*$. But in the motion of a conservative system, its total energy retains the initial magnitude E_0 and consequently $E < E^*$ during the whole time of motion. For this reason, when a system is in motion, the point depicting this motion in state space cannot attain the boundary of the ε-neighbourhood on which $E \geqslant E^*$ and all the time lies inside this neighbourhood.

The theorem is proved.

Note two things with regard to this theorem.

Note 1. The Lagrange theorem remains valid for a nonconservative system that is obtained from a conservative system by the addition of gyroscopic and dissipative forces. Indeed, first of all note that the equilibrium position is preserved if gyroscopic or dissipative forces $\tilde{Q}_\nu(q_k, \dot{q}_k)$ are added to the system. For these forces

$$\sum_{\nu=1}^{n} \tilde{Q}_\nu (q_k, \dot{q}_k) \dot{q}_\nu \leqslant 0, \quad |q_k| < \Delta, \quad |\dot{q}_k| < \Delta$$

As before, choose for the equilibrium position the coordinate origin of the state space and assume that among the functions $\tilde{Q}_\nu$ there is at least one $\tilde{Q}_j$ such that $\tilde{Q}_j(0) \neq 0$. Then, due to continuity, $\tilde{Q}_j \neq 0$ in the Δ-neighbourhood of the coordinate origin as well. But since q_k and $\dot{q}_k$ are independent, their values in this neighbourhood may always be chosen so that

$$\sum_{\nu=1}^{n} \tilde{Q}_\nu (q_k, \dot{q}_k) \dot{q}_\nu > 0$$

But this contradicts the hypothesis of dissipativity or gyroscopicity of the forces. Therefore, the supposition about the existence of $\tilde{Q}_j(0) \neq 0$ leads to a contradiction, i.e. all $\tilde{Q}_\nu(0) = 0$, $\nu = 1, \ldots, n$, and this indicates that the addition of gyroscopic and dissipative forces does not violate the equilibrium.

Gyroscopic forces do not upset the law of conservation of total energy (see Sec. 8) and for this reason the whole proof of the Lagrange theorem remains unchanged in the case of gyroscopic forces as well. In the case of dissipative forces, the total energy $E = T + \Pi$ diminishes as the system moves, and consequently when the system is in motion we have the inequality $E \leqslant E_0$ in place of the equality $E = E_0$. But from this it also follows that throughout the motion $E < E^*$, if $E_0 < E^*$. Thus the proof of the theorem is preserved here too with that slight modification.

Note 2. The equilibrium position of a conservative system will also be stable when the potential energy Π in this position has a *nonstrict* minimum, but in any ε-neighbourhood of the equilibrium position there exists a closed hypersurface

$$f(q_1, \ldots, q_n) = 0 \tag{9}$$

that contains the equilibrium position within it and has the property that on this hypersurface the values of the potential energy are strictly greater than the value of Π in the equilibrium position.

Indeed, as before, in the equilibrium position let $q_1 = \ldots = q_n = 0$ and $\Pi(0, \ldots, 0) = 0$. Also, let the equation of the hypersurface (9) be chosen so that for points located inside the closed hypersurface (9) the following inequality holds:

$$f(q_1, \ldots, q_n) > 0 \tag{10}$$

Then this inequality, together with the inequalities

$$|\dot{q}_k| < c_k \quad (0 < c_k < \varepsilon;\ k = 1, \ldots, n) \tag{11}$$

defines, in the $2n$-dimensional state space, a region (a finite "hypercylinder") G located inside the ε-neighbourhood (6). On the boundary of the region G, either $f = 0$ (then $\Pi > 0$, $T \geqslant 0$) or at least for one k the equality $|\dot{q}_k| = c_k$ holds (then $T > 0$, $\Pi \geqslant 0$). Therefore, the strict inequality $E = T + \Pi > 0$ always holds on the boundary of the region G.

In the case under consideration, the minimum of the function E on the boundary of the ε-neighbourhood (6) can be zero. Then when proving the Lagrange theorem, in place of the ε-neighbourhood one has to take the region G situated inside it. On the boundary of the region G the minimum of total energy $E^* > 0$. From then on the remainder of the proof remains without change.

For $n = 1$ the closed hypersurface (9) degenerates into a collection of two points on the q-axis located on different sides of the origin O, while the region G degenerates into a rectangle situated inside the ε-neighbourhood of the point O (Fig. 43).

The reasoning given in Note 2 holds also for the case when gyroscopic and dissipative forces (see Note 1) are additionally applied to the system.

If the points at which the function Π has a minimum $\Pi = 0$ fill the solid curve emanating from the equilibrium position, then this equilibrium position can also be unstable. To illustrate, take a free particle with potential energy that does not contain one of the coordinates, for example x: $\Pi = \Pi(y, z)$, and $\Pi(0, 0) = 0$ and $\Pi(y, z) > 0$ for $y^2 + z^2 > 0$. In this case, the minimum points fill the x-axis. The equilibrium position $x = y = z = 0$ is unstable, since for an arbitrarily small (in magnitude) initial velocity along the x-axis, the point will execute uniform motion along the x-axis.

In the Examples 1 and 2 given on page 167, a conservative system was considered, and in Example 3 (page 167) a dissipative force acts on a particle as well. The potential energy has a strict minimum

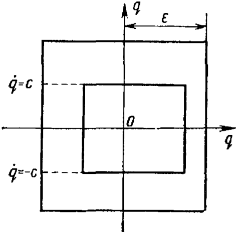

Fig. 43

in Example 1 at the lowest point of the neighbourhood, and in Examples 2 and 3 at $x = 0$. For this reason, these equilibrium positions are called stable.

Example 4. A conservative system with one degree of freedom has a potential energy $\Pi = q^4 \sin^2 \frac{1}{q}$ [we also define $\Pi(0) = 0$]. In accordance with Note 2, the position $q = 0$ is a stable position of equilibrium.

34. Criteria of Instability of an Equilibrium Position. Theorems of Lyapunov and Chetayev

As far back as 1892, A. M. Lyapunov, in his celebrated dissertation entitled "The General Problem of Stability of Motion", posed the question of the converse of the Lagrange theorem. The problem has still not been solved completely. A partial solution is given by two theorems of Lyapunov and a theorem of Chetayev, in which certain sufficient conditions are established for instability of an equilibrium position.

As before, let $q_1 = \ldots = q_n = 0$ and $\Pi(0, \ldots, 0) = 0$ in an equilibrium position. Write an expansion of the potential energy in a power series of coordinates ("deviations"):

$$\Pi = \Pi_m(q_1, \ldots, q_n) + \Pi_{m+1}(q_1, \ldots, q_n) + \ldots$$
$$[\Pi_m(q_1, \ldots, q_n) \not\equiv 0, \ m \geqslant 2] \qquad (1)$$

where $\Pi_k (q_1, \ldots, q_n)$ is a homogeneous function of the kth degree ($k = m, m+1, \ldots$), and the lowest of the powers of terms actually appearing in the expansion is $m \geqslant 2$, since all $\left(\frac{\partial \Pi}{\partial q_i}\right)_0 = 0$ in the position of equilibrium.

Lyapunov's first theorem. *If the potential energy* $\Pi (q_1, \ldots, q_n)$ *of a conservative system in a position of equilibrium does not have a minimum and this circumstance may be seen from the second-degree terms* $\Pi_2 (q_1, \ldots, q_n)$ *of the expansion* (1)*, *then the given position of equilibrium is unstable.*

Proof. In the expression for the kinetic energy $T = \frac{1}{2} \sum_{i, k=1}^{n} a_{ik} \dot{q}_i \dot{q}_k$, expand the coefficients $a_{ik} (q_1, \ldots, q_n)$ in a power series of coordinates and denote by a_{ik}^0 the absolute terms (i.e. the values of the functions $a_{ik} (q_1, \ldots, q_n)$ for $q_1 = \ldots = q_n = 0$). Then, putting $T_0 = \frac{1}{2} \sum_{i, k=1}^{n} a_{ik}^0 \dot{q}_i \dot{q}_k$, we will have

$$T = T_0 + (*), \quad \Pi = \Pi_2 + (*)$$

here and henceforward we shall use the symbol $(*)$ to designate the sum of the terms having a higher order of smallness relative to the coordinates and velocities than the terms written out earlier. Since T_0 is a positive definite quadratic form with constant coefficients, it is possible, by means of a nonsingular linear transformation of the variables, to simultaneously reduce the two quadratic forms T_0 and Π_2 to a sum of squares, after which the expansion for T and Π in the new variables $\theta_1, \ldots, \theta_n$ will take the form**

$$T = \frac{1}{2} \sum_{k=1}^{n} \dot{\theta}_k^2 + (*), \quad \Pi = \frac{1}{2} \sum_{k=1}^{n} \lambda_k \theta_k^2 + (*) \tag{2}$$

Since the quadratic form Π_2 assumes certain negative values, at least one $\lambda_k < 0$.

The Lagrange equations may be written in the coordinates θ_k thus

$$\ddot{\theta}_k = -\lambda_k \theta_k + (*) \quad (k = 1, \ldots, n) \tag{3}$$

---

* That is, $m = 2$, and Π_2 is a quadratic form assuming negative values (perhaps along with positive values).

** The coordinates $\theta_1, \ldots, \theta_n$ are called *normal* or *principal* coordinates. They will be considered in more detail in Sec. 41.

Consider the auxiliary quadratic form

$$V = -\frac{1}{2}\sum_{k=1}^{n} \varepsilon_k \left[\left(\lambda_k^2 + \lambda_k + \frac{\mu^2}{2}\right)\theta_k^2 + \mu(1-\lambda_k)\theta_k\dot{\theta}_k + \right.$$
$$\left. + \left(1 + \lambda_k + \frac{\mu^2}{2}\right)\dot{\theta}_k^2 \right] \tag{4}$$

where $\varepsilon_k = 1$ for $\lambda_k \geqslant 0$ and $\varepsilon_k = -1$ for $\lambda_k < 0$ $(k = 1, \ldots, n)$, and the number $\mu > 0$.

It is verified directly that, by virtue of equations (3) and equality (4),

$$\frac{d}{dt}(e^{-\mu t}V) = e^{-\mu t}\left[\mu \sum_{k=1}^{n} \varepsilon_k \left(\lambda_k + \frac{\mu^2}{4}\right)(\theta_k^2 + \dot{\theta}_k^2) + (*) \right] \tag{5}$$

when a system is in motion.

Without loss of generality, we assume that $\lambda_1 < 0$, and that λ_1 is the largest of the negative numbers λ_k. The positive number μ is chosen so that the following inequalities hold simultaneously:

$$\lambda_1 + \frac{\mu^2}{4} < 0, \quad \lambda_1^2 + \lambda_1 + \frac{\mu^2}{2} > 0 \tag{6}$$

From the first inequality it follows that the sum of the right-hand side of (5) is a positive definite quadratic form. But then, for θ_k and $\dot{\theta}_k$ sufficiently small (in absolute value),

$$|\theta_k| < \Delta, \quad |\dot{\theta}_k| < \Delta \quad (k = 1, \ldots, n) \tag{7}$$

the right-hand side of (5) will always be positive, i.e.

$$\frac{d}{dt}(e^{-\mu t}V) > 0$$

whence

$$e^{-\mu t}V > e^{-\mu t_0}V_0$$

or

$$V > V_0 e^{\mu(t-t_0)} \tag{8}$$

Put all the initial values θ_k^0, $\dot{\theta}_k^0$ $(k = 1, \ldots, n)$ equal to zero, with the exception of θ_1^0, which we shall take, in absolute value, less than Δ. Then, using expression (4) and the second inequality of (6), we find $V_0 > 0$. But in that case the motion will definitely go beyond the limits of the neighbourhood of (7), however small $|\theta_1^0|$ may be, since otherwise it would follow from inequality (8) that $\lim\limits_{t\to\infty} V = \infty$, while the quadratic form V in the neighbourhood of (7) is bounded.

The theorem is proved.

For the case when $m > 2$, in expansion (1) we can use the following two theorems that are given without proof.*

Lyapunov's second theorem. *If the potential energy* Π *of a conservative system for* $q_1 = \ldots = q_n = 0$ *has a strict maximum and this circumstance may be determined by proceeding from the terms of lowest power of* $\Pi_m (q_1, \ldots, q_n)$ $(m \geqslant 2)$ *in the expansion* (1),** *then the position* $q_1 = \ldots = q_n = 0$ *is an unstable position of equilibrium of the system.*

Chetayev's theorem. *If the potential energy* Π *of a conservative system is a homogeneous function of the deviations* $q_1, \ldots, q_n$ *and in the position of equilibrium* $q_1 = \ldots = q_n = 0$ *does not have a minimum, then this equilibrium position is unstable.*

Example 1. Let $\Pi = A(1 - \cos \alpha q)$; $n = 1$. The function Π has strict minima at the points $q_{2k} = \frac{2k\pi}{\alpha}$ and strict maxima ($k = 0, \pm 1, \pm 2, \ldots$) at the points $q_{2k-1} = \frac{(2k-1)\pi}{\alpha}$. The last circumstance follows from the form of the lowest-order term in the expansion in powers of the deviations

$$\Pi = -\frac{\alpha^2}{2}(q - q_{2k-1})^2 + \ldots$$

Then, according to the theorems of Lagrange and Lyapunov, to the points q_{2k} there correspond stable equilibrium positions, and to the points q_{2k-1}, unstable ones.

Example 2. $\Pi = Aq_1 \ldots q_n$. From Chetayev's theorem it follows that the position $q_1 = \ldots = q_n = 0$ is an unstable position of equilibrium.

35. Asymptotic Stability of an Equilibrium Position. Dissipative Systems

Now introduce the concept of an asymptotically stable position of equilibrium. An equilibrium position is termed asymptotically stable if it is stable and if, besides, for sufficiently small (in absolute value) initial deviations and initial velocities, all deviations and velocities tend to zero as the time t increases without bound, i.e. if there exists a number $\delta_0 > 0$ such that

$$\lim_{t\to\infty} q_i(t) = 0, \quad \lim_{t\to\infty} \dot{q}_i(t) = 0 \qquad (i = 1, \ldots, n) \tag{1}$$

every time the inequalities

$$|q_i^0| < \delta_0, \quad |\dot{q}_i^0| < \delta_0 \quad (i = 1, \ldots, n) \tag{1'}$$

---

* The reader will find the proofs in the books by Lyapunov [16], Chetayev [5], and Malkin [18].

** That is, in a certain neighbourhood of the origin (excluding the origin itself) always $\Pi_m(q_1, \ldots, q_n) < 0$. This is only possible if m is even.

are satisfied. In a geometrical interpretation (Fig. 44), this signifies that in the state space $(q_i, \dot{q}_i)$ all trajectories emanating from the δ_0-neighbourhood of the origin O asymptotically approach the point O as $t \to \infty$.

Of the Examples 1, 2 and 3 examined on pages 167-168, only in Example 3 is the stable equilibrium position asymptotically stable.

We shall consider scleronomic systems under the action of the potential forces $-\frac{\partial \Pi}{\partial q_i}$ and the nonpotential forces $\tilde{Q}_i$ $(i = 1, \ldots, n)$

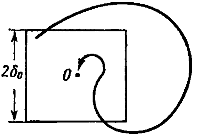

Fig. 44

assuming that the potential energy Π and the nonpotential forces $\tilde{Q}_i$ do not depend explicitly on the time:

$$\Pi = \Pi(q_k), \quad \tilde{Q}_i = \tilde{Q}_i(q_k, \dot{q}_k) \quad (i = 1, \ldots, n) \tag{2}$$

In this case the time t does not enter explicitly into the Lagrange equations, which may be written in the following form (solved for the generalized accelerations) (see Sec. 7, page 47):

$$\ddot{q}_i = G_i(q_k, \dot{q}_k) \quad (i = 1, \ldots, n) \tag{3}$$

In the case under consideration, the total energy E of the scleronomic system does not contain the time explicitly:

$$E = E(q_k, \dot{q}_k) \tag{4}$$

Computing its total derivative with respect to the time when the system is in motion, we find

$$\frac{dE}{dt} = \sum_{i=1}^{n} \frac{\partial E}{\partial q_i} \dot{q}_i + \frac{\partial E}{\partial \dot{q}_i} G_i = E'(q_k, \dot{q}_k) \tag{5}$$

Thus, at every point of the state space $(q_k, \dot{q}_k)$, not only the total energy but also its total derivative with respect to the time has a definite value.

If the forces $\tilde{Q}_i$ $(i=1, \ldots, n)$ are dissipative forces (see Sec. 8), then $\frac{dE}{dt} \leqslant 0$ when the system is in motion, i.e. in the region of state space under consideration the function $E'(q_k, \dot{q}_k)$ assumes only nonpositive values.

In the case of a definite dissipative system (see Sec. 8), $\frac{dE}{dt} = E'(q_i, \dot{q}_i)$ vanishes only at those points of the state space $(q_i, \dot{q}_i)$ where all $\dot{q}_i = 0$, $i = 1, \ldots, n$. We shall assume that the equilibrium position of the system is isolated, i.e. there are no other equilibrium positions in its neighbourhood. Then we have the following:

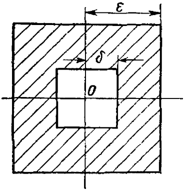

Fig. 45

Theorem on asymptotic stability. *If the potential energy* Π *of a scleronomic definite dissipative system has a strict minimum in some equilibrium position that is isolated, then the equilibrium position is asymptotically stable.*

Proof. Again let

$$q_1 = q_2 = \ldots = q_n = 0 \quad \text{and} \quad \Pi(0) = 0$$

in the equilibrium position. As in the proof of the Lagrange theorem, choose in the state space an ε-neighbourhood of the coordinate origin O in which the energy E is positive,

$$E(q_i, \dot{q}_i) > 0 \qquad (6)$$

at all points different from O and in which there are no equilibrium states different from O.

Since according to the Lagrange theorem, the position of equilibrium $q_1 = \ldots = q_n = 0$ is stable, for any $\varepsilon > 0$ it is possible to indicate a $\delta(\varepsilon) > 0$ such that all motions occur inside the ε-neighbourhood of the point O if the initial point is chosen in the δ-neighbourhood (Fig. 45). For the δ_0-neighbourhood take the neighbourhood in which condition (8) of Sec. 33 ($\delta_0 \leqslant \delta$) is fulfilled. We consider one of the motions initiated in the δ_0-neighbourhood. Since, given motion, the energy E diminishes, it follows* that

$$\lim_{t\to\infty} E(t) = E_\infty \geqslant 0$$

and $E(t) \geqslant E_\infty$ ($t \geqslant t_0$). First assume that $E_\infty \neq 0$, i.e. $E_\infty > 0$, and consider the sequence of the values of the time $t_s \longrightarrow \infty$ and the sequence of values of the phase coordinates

$$q_i^{(s)} = q_i(t_s),\quad \dot{q}_i^{(s)} = \dot{q}_i(t_s),\quad i = 1, \ldots, n$$

Since the entire trajectory lies in the ε-neighbourhood, the following inequalities hold for all s and i:

$$|q_i^{(s)}| < \varepsilon,\quad |\dot{q}_i^{(s)}| < \varepsilon$$

By virtue of the Bolzano-Weierstrass lemma, it is possible, from the infinite bounded sequences $q_i^{(s)}$ and $\dot{q}_i^{(s)}$, to choose the subsequences $q_i^{(k)}$ and $\dot{q}_i^{(k)}$. Let

$$\lim_{k\to\infty} q_i^{(k)} = q_i^*,\quad \lim_{k\to\infty} \dot{q}_i^{(k)} = \dot{q}_i^*,\quad i = 1, \ldots, n \tag{7}$$

where

$$|q_i^*| < \varepsilon,\quad |\dot{q}_i^*| < \varepsilon,\quad i = 1, \ldots, n$$

But then, because $E(t)$ is continuous,

$$E(q_i^*, \dot{q}_i^*) = \lim_{k\to\infty} E(q_i^{(k)}, \dot{q}_i^{(k)}) = E_\infty > 0$$

By hypothesis, the point $(q_i^*, \dot{q}_i^*)$ does not coincide with the coordinate origin, where $E = 0$.

Let us take the point $(q_i^*, \dot{q}_i^*)$, $i = 1, \ldots, n$ for the initial point of motion when $t = t_0$. Since this point does not coincide with the point O, i.e. it is not an equilibrium position, then when the system is in motion at least one of the generalized velocities $\dot{q}_i$ will be different from zero and for this reason $\frac{dE}{dt} < 0$. But then for some $t = t_1$ the inequality $E < E_\infty$ will hold.

Further, let us consider the motion of a system starting at $t = t_0$ from the point $(q_i^{(k)}, \dot{q}_i^{(k)})$. By virtue of (7), the values of $q_i^{(k)}$

---

[*] This limit exists by virtue of the fact that $E(t)$ is a continuous monotonic decreasing nonnegative function.

and $\dot{q}_i^{(k)}$, for sufficiently large k, will be arbitrarily close to the values of q_i^* and $\dot{q}_i^*$ respectively. Consequently, for sufficiently large k, the values of the phase coordinates for $t = t_1$ will also be arbitrarily close in the case of motions starting, when $t = t_0$, from the initial points $(q_i^{(k)}, \dot{q}_i^{(k)})$ and $(q_i^*, \dot{q}_i^*)$ (the solutions of the systems of differential equations are continuous functions of initial data). Therefore, at $t = t_1$ the inequality $E(t_1) < E_\infty$ will hold for a motion that started, when $t = t_0$, from the point $(q_i^{(k)}, \dot{q}_i^{(k)})$ since the total energy E is a continuous function of the phase coordinates. But by virtue of the uniqueness of the solutions of the Lagrange equations, the state (at time $t = t_1$) of a system that at $t = t_0$ started from an initial point $(q_i^{(k)}, \dot{q}_i^{(k)})$ $i = 1, \ldots, n$, coincides* with the state at time $t_k + t_1$ of a system which at $t = t_0$ started from the initial point $(q_i^{(0)}, \dot{q}_i^{(0)})$, $i = 1, \ldots, n$; for this reason, the value of the energy $E(t_k + t_1)$ that interests us must satisfy the inequality

$$E(t_k + t_1) < E_\infty$$

but this is impossible, since $E(t) \geqslant E_\infty$ for any $t > t_0$.

We have thus arrived at a contradiction by assuming that $E_\infty \neq 0$; hence,

$$E_\infty = 0 \quad \text{and} \quad \lim_{t \to \infty} E(t) = 0 \tag{8}$$

Since, by virtue of the inequality (6), the equality $E = 0$ occurs only at the point O, it follows from (8) that as $t \to \infty$ the point depicting the system in state space tends to the coordinate origin, which means that the relations (1) hold. The theorem is proved.

In Example 3 on page 167, $E = \frac{1}{2} m\dot{x}^2 + \frac{1}{2} cx^2$ and

$$\frac{dE}{dt} = m\dot{x}\ddot{x} + cx\dot{x} = (m\ddot{x} + cx)\,\dot{x} = -2f\dot{x}^2 < 0$$

since $f > 0$. The system is definite dissipative and the equilibrium position is isolated; this follows from the equation of motion when substituting the solution $x = \text{const}$. By virtue of the theorem, the position of equilibrium is asymptotically stable.

---

* Since the time t does not appear explicitly in the equation of motion (3), choice of zero time is immaterial. Therefore, if in place of $q_i^{(0)}$, $\dot{q}_i^{(0)}$ we take for the initial state $q_i^{(k)}$, $\dot{q}_i^{(k)}$, $i = 1, \ldots, n$, then the system will subsequently pass through the same states as in the initial motion, but at different instants of time.

An investigation of the motion of a scleronomic definite dissipative system in the neighbourhood of its asymptotically stable equilibrium position will be given in Sec. 46.

Lyapunov proved a theorem that generalizes the Lagrange theorem. He noticed that in proving the Lagrange theorem it is possible, in place of the energy E, to take any continuous (with continuous first-order partial derivatives) function $V(q_k, \dot{q}_k)$ which has a strict minimum in the equilibrium state and which does not increase under any motion of the system.

Compute the time derivative of the function V by taking advantage of the equation of motion (3):

$$\frac{dV}{dt} = \sum_{i=1}^{n} \frac{\partial V}{\partial q_i} \dot{q}_i + \frac{\partial V}{\partial \dot{q}_i} G_i(q_k, \dot{q}_k) = V'(q_k, \dot{q}_k) \tag{9}$$

One thing that follows from this formula is that the function V' vanishes at the coordinate origin O of the state space, since the point O corresponds to the state of equilibrium in which all the $\dot{q}_i = 0$ and all the $\ddot{q}_i = G_i = 0$. If the function V does not increase in any motion, the $\frac{dV}{dt} = V'(q_k, \dot{q}_k) \leqslant 0$. In this case, the function V' has a maximum in the state of equilibrium O. Now if this maximum is strict, then in the neighbourhood of the point O (with the exception of the point O itself) $V' < 0$, and when the system is in motion within the limits of this neighbourhood, the function V decreases strictly.

It is now possible to repeat almost literally the proof of the Lagrange theorem, using the function V in place of E. In the case of asymptotic stability (for example, for a dissipative system), the proof will even be simpler if we demand that in the equilibrium position the derivative $\frac{dV}{dt}$ have a strict extremum of a type opposite to the extremum of the function V.*

Note also that in formulating the stability criterion, the words "minimum" and "maximum" can be interchanged, since substituting the function $-V$ for the function V returns us to the earlier formulation.

Thus the following theorem has been proved:

Theorem. If there is given the equilibrium position of a scleronomic system that is under the action of forces which do not depend explicitly on the time, and there exists a function $V(q_k, \dot{q}_k)$ that is continuous together with its first partial derivatives and that has a strict extremum in the given state of equilibrium, whereas the derivative V' of V with respect to the time (computed by the equations of motion) has in this same state an extremum of the opposite type, then the equilibrium position under consideration is stable. If the extremum of the derivative is also strict, then the equilibrium position is asymptotically stable.

---

* In the proof of the theorem on the stability of a dissipative system, the derivative $\frac{dE}{dt}$ did not have a strict maximum in the equilibrium position. In this connection we had to stipulate that the equilibrium position was isolated.

The function $V(q_k, \dot{q}_k)$ spoken of in the theorem is generally called the *Lyapunov function.*

36. Conditional Stability. General Statement of the Problem. Stability of Motion or of an Arbitrary Process. Lyapunov's Theorem

Inequalities (1) and (2) (see page 166) that define the stability of an equilibrium position involved all the deviations q_i and all the generalized velocities $\dot{q}_i$. However, in many problems we encounter a conditional stability when the indicated inequalities hold for certain of the $2n$ quantities $q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n$ or, in a more general statement, for certain functions $x_1, \ldots, x_m$ of these quantities:

$$x_i = f_i(q_k, \dot{q}_k) \quad (i = 1, \ldots, m) \tag{1}$$

It is also assumed that all the functions (1) vanish for

$$q_k = 0, \ \dot{q}_k = 0 \ (k = 1, \ldots, n), \text{ i.e. } f_i(0, \ldots, 0) = 0$$

and satisfy the autonomous* system of ordinary first-order differential equations**

$$\frac{dx_i}{dt} = X_i(x_1, \ldots, x_m, t) \ (i = 1, \ldots, m) \tag{2}$$

where $X_i(x_1, \ldots, x_m, t)$ $(i = 1, \ldots, m)$ are continuous functions in the region

$$|x_i| \leqslant \Delta, \ t \geqslant t_0 \ (i = 1, \ldots, m) \tag{3}$$

(t_0 is a fixed initial instant of time).

To the equilibrium state there corresponds a zero solution $x_i \equiv 0$ $(i = 1, \ldots, m)$ of the system of differential equations (2).

---

* See footnote on page 81.

** In investigating the conditional stability of scleronomic systems, the functions X_i $(i = 1, \ldots, m)$ do not depend explicitly on t. We have put t on the right sides as arguments, having in view nonscleronomic systems and subsequent generalizations.

Such a solution presupposes that the right-hand sides of equations (2) satisfy the condition

$$X_i(0, \ldots, 0, t) \equiv 0 \quad (i = 1, \ldots, m) \tag{4}$$

From the mathematical point of view it is a question of the stability of a zero solution of the system of differential equations (2), this stability being defined as follows: for any $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon) > 0$ such that

for any $t \geqslant t_0$

$$|x_i(t)| < \varepsilon \quad (i = 1, \ldots, m) \tag{5}$$

as long as

$$|x_i(t_0)| < \delta \quad (i = 1, \ldots, m) \tag{6}$$

For a geometric interpretation of the inequalities (5) and (6), use is made of the ε- and δ-neighbourhoods of the coordinate origin in the m-dimensional space $(x_1, \ldots, x_m)$. In the case of asymptotic stability we also require the existence of a $\delta_0 > 0$ such that

$$\lim_{t \to \infty} x_i(t) = 0 \quad (i = 1, \ldots, m) \tag{7}$$

so long as

$$|x_i(t_0)| < \delta_0 \quad (i = 1, \ldots, m) \tag{8}$$

If the stability (not conditional!) of an equilibrium position is under study, then for $x_1, \ldots, x_m$ we can take the quantities $q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n$ or $q_1, \ldots, q_n, p_1, \ldots, p_n$. In the former case, the equations (2) are the Lagrange equations written as a system of $2n$ first-order differential equations in the unknown functions $q_1, \ldots, \dot{q}_n$. In the latter case, equations (2) are the canonical equations of Hamilton:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad (i = 1, \ldots, n) \tag{9}$$

Let us consider two important special cases of the system of equations (2) which are frequently encountered in applications.

1°. Stationary case, when t does not appear explicitly on the right-hand sides X_i of equations (2), that is, when

$$\frac{\partial X_i}{\partial t} = 0 \quad (i = 1, \ldots, m)$$

2°. Periodic case, when the right-hand sides X_i have a period τ with respect to the variable t:

$$X_i(x_1, \ldots, x_m, t + \tau) = X_i(x_1, \ldots, x_m, t) \quad (i = 1, \ldots, m)$$

In both cases the stability of the zero solution of the system of differential equations (2) is determined by a theorem that is a direct generalization of the theorem given at the end of Sec. 35.

Lyapunov's Theorem. *If in the stationary or the periodic case there exists a function* $V(x_1, \ldots, x_m, t)$ *continuous together with its first partial derivatives in the region* (3), *which function, for any* t *considered as a parameter, has at the point* $x_1 = \ldots = x_m = 0$ *a strict extremum, whereas again at the same point, and for any* t, *its time derivative* $V'(x_1, \ldots, x_m, t) = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} X_i + \frac{\partial V}{\partial t}$ *has an extremum of opposite type, then the zero solution of the system* (2) *is stable. Now if the extremum of the derivative* V' *is also strict, then the zero solution of the system* (2) *is asymptotically stable. Here it is assumed that in the stationary case the function* V *is not explicitly dependent on the time* t, *but in the periodic case this function is periodic with respect to* t *with period* τ [τ is the period of the right-hand sides of equations (2)].

It suffices to prove this theorem for the periodic case since the stationary case may be regarded as a special case of the periodic with any τ. The proof consists in repeating the reasoning given earlier in the proof of the Lagrange theorem and the theorem on asymptotic stability with the following modifications: in place of E we now use the difference

$$V(x_1, \ldots, x_m, t) - V(0, \ldots, 0, t)$$

and in place of the space $(q_i, \dot{q}_i)$ we take the m-dimensional space $(x_1, \ldots, x_m)$. We regard t, which appears in V, as a parameter. This parameter, by virtue of its periodicity, may be varied in the finite interval $t_0 \leqslant t \leqslant t_0 + \tau$. Because of this circumstance, the presence of the variable t in V does not give rise to any complications in the proof of the theorem.

Note that in the general (nonstationary and nonperiodic) case, more rigorous conditions [see 5] must be imposed on the Lyapunov function.

We note one special case of the Lyapunov theorem, which is frequently used as a test for simple (nonasymptotic) stability.

Let the function $V(x_1, \ldots, x_m, t)$ be the integral of the system of differential equations (2), i.e. the function V becomes a constant upon substitution into it of any solution of the system (2). In this case $\frac{dV}{dt} = V'(x_1, \ldots, x_m, t) \equiv 0$, and it may be taken that the function V' has a maximum and minimum (nonstrict, of course) at the point $x_1 = \ldots = x_m = 0$ for any t. For this reason, we have the following corollary to the Lyapunov theorem:

Corollary. If the system of differential equations (2) has an integral $V(x_1, \ldots, x_m, t)$ [which is not dependent on t in the stationary case and is periodic with respect to t with period τ in the periodic case] and this integral has a strict extremum at the point $x_1 = \ldots = x_m = 0$ for any fixed t, then the zero solution of the system (2) is stable.

Note that in proving the Lagrange theorem for a conservative system, use is made of the energy integral E.

Now let us consider motions or more general processes described by the system of first-order differential equations

$$\frac{dz_i}{dt} = Z_i(z_i, \ldots, z_m, t) \quad (i = 1, \ldots, m) \tag{10}$$

where the right-hand sides are continuous functions in some region of variation of the variables $z_1, \ldots, z_m$ for $t \geqslant t_0$ that satisfy the conditions of existence and uniqueness of solution on the basis of the initially given $z_i(t_0)$ $(i = 1, \ldots, m)$.

Let $\tilde{z}_i(t)$ $(i = 1, \ldots, m)$ be a solution of the system of equations (10) defining some process. To elucidate the question of the stability of this process, we introduce new unknown functions or "deviations", in place of the unknown functions $z_1, \ldots, z_m$:

$$x_i = z_i - \tilde{z}_i(t) \quad (i = 1, \ldots, m) \tag{11}$$

Then in the new variables, the system of differential equations (10) will be written in the form of the system (2), where

$$X_i = Z_i[x_1 + \tilde{z}_1(t), \ldots, x_m + \tilde{z}_m(t)] - X_i[\tilde{z}_1(t), \ldots, \tilde{z}_m(t)] \tag{12}$$

The zero solution $x_1 = \ldots = x_m = 0$ of the system of differential equations (2) corresponds to the solution $z_i = \tilde{z}_i(t)$ $(i = 1, \ldots, m)$ of the system of differential equations (10) in the new variables. This circumstance permits the question of the stability of the process $\tilde{z}_i(t)$ $(i = 1, \ldots, m)$ to be reduced to the problem, which we have already studied, of the stability of the zero solution of the system of differential equations (2). In other words, the solution $z_i = \tilde{z}_i(t)$ $(i = 1, \ldots, m)$ of the system (10) is called stable (or correspondingly, asymptotically stable) if the zero solution $x_1 = \ldots = x_m = 0$ of the system of equations in deviations (2) with the right-hand sides defined by the formulas (12), is stable (or correspondingly, asymptotically stable). All this opens up a wide field for application of the Lyapunov theorem given in this section. This theorem may be used to determine not only the stability of an equilibrium position, but also the stability of motion, in general, of any process defined by a system of ordinary differential equations.

Example 1. Consider the inertial rotation of a solid having a fixed point O. Euler's dynamical equations in this case have the form

$$A\frac{dp}{dt}=(B-C)\,qr,\quad B\frac{dq}{dt}=(C-A)\,rp,\quad C\frac{dr}{dt}=(A-B)\,pq \tag{13}$$

Here, p, q, r are projections of the angular velocity ω on the principal axes of inertia of the solid, $O\xi$, $O\eta$, $O\zeta$; and A, B, C are the moments of inertia about these axes.

Equations (13) allow for the following three particular solutions defining permanent rotations of the solid about the principal axes:

$$\begin{aligned}&1^\circ\ q=r=0,\quad p=\text{const}=p_0\\ &2^\circ\ r=p=0,\quad q=\text{const}=q_0\\ &3^\circ\ p=q=0,\quad r=\text{const}=r_0\end{aligned}$$

We confine ourselves to working out the stability of rotation 1°, since 2° and 3° may be written as 1° with the axes labelled differently. Here, the stability of the solution 1° of Euler's equations (13) will determine the conditional stability of rotation of 1° with respect to the angular velocity ω.*

Form the equations in the deviations, putting $x_1 = p - p_0$, $x_2 = q$, $x_3 = r$:

$$\left.\begin{aligned}&\frac{dx_1}{dt}=\frac{B-C}{A}x_2x_3,\quad \frac{dx_2}{dt}=\frac{C-A}{B}x_3(x_1+p_0)\\ &\frac{dx_3}{dt}=\frac{A-B}{C}x_2(x_1+p_0)\end{aligned}\right\} \tag{14}$$

Let the major or minor axis of the ellipsoid of inertia be along the axis $O\xi$ of the permanent rotation under study. The fact that A, B, and C are inversely proportional to the squares of the axes of the ellipsoid of inertia means that $A < B$. C or $A > B$, C. For the Lyapunov function, let us take the function

$$V=(Ax_1^2+Bx_2^2+Cx_3^2+2Ap_0x_1)^2\pm[B(B-A)\,x_2^2+C(C-A)\,x_3^2]$$

where the plus sign is taken for $A < B$, C and the minus sign for $A > B$, C.

The function V vanishes at $x_1 = x_2 = x_3 = 0$ and is positive in the neighbourhood of this point, that is, the function V has a strict minimum at this point. On the other hand, it is readily verified that $\frac{dV}{dt} = 0$ by virtue of the equations (14), i.e. the function V is an integral of the system of differential equations (14). For this reason, in accord with the corollary to the Lyapunov theorem, permanent rotation about the major or minor axis of an ellipsoid of inertia is stable.

It is also possible to show that permanent rotation about the mean axis of an ellipsoid of inertia is unstable, but this would require use of the instability test of Chetayev [5].

Example 2. By way of illustration, consider the problem of stability of rotational motion of an artillery shell.

To simplify the problem, assume that the centre of gravity of the shell C is in rectilinear motion along the horizontal x-axis with constant velocity $v=\text{const}$. Let Cxy be the vertical plane of firing. The position of the axis $C\xi$

---

* Stability with respect to ω means that a small change in the initial angular velocity ω_0 involves a small change in the vector ω throughout the motion. In other words, this is stability with respect to deviations $x_1 = p - p_0$, $x_2 = q$, $x_3 = r$. Naturally Euler's angles steadily increase here.

(the axis of dynamical symmetry) of the shell is defined by two angles: angle α formed by projection Cx_1 of the axis $C\xi$ on the Cxy-plane with axis Cx, and angle β between Cx_1 and $C\xi$ (Fig. 46). By successive rotations through the angle α and through the angle β, the trihedron of axes $Cxyz$ passes into the trihedron $C\xi\eta\zeta$. A supplementary rotation through the angle φ about the axis $C\xi$ carries

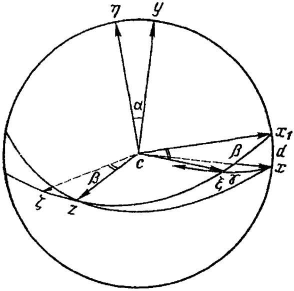

Fig. 46

the trihedron $C\xi\eta\zeta$ into the trihedron of axes fixed in the shell. Therefore, the angular velocity $\boldsymbol{\omega}$ of the shell consists of three components:

$$\boldsymbol{\omega} = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_2 + \boldsymbol{\omega}_3$$

where $\omega_1 = \dot{\alpha}$, $\omega_2 = \dot{\beta}$ and $\omega_3 = \dot{\varphi}$. The projections of the angular velocity on the principal axes of inertia $C\xi$, $C\eta$, $C\zeta$ are defined by the formulas

$$p = \dot{\varphi} + \dot{\alpha} \sin\beta$$

$$q = -\dot{\beta}, \quad r = \dot{\alpha} \cos\beta$$

Denoting by A the axial and by B the equatorial moment of inertia of the shell, we get the following expression for the kinetic energy:

$$T = \frac{1}{2}\left[Ap^2 + B\left(q^2 + r^2\right)\right] = \frac{1}{2}\left[A\left(\dot{\varphi} + \dot{\alpha}\sin\beta\right)^2 + B\left(\dot{\beta}^2 + \dot{\alpha}^2\cos^2\beta\right)\right]$$

We shall assume that besides the force of gravity there is applied, at point D on the axis of the shell (in the "centre of pressure"), the force of air drag $\boldsymbol{R}$, which is constant* and directed opposite to the velocity $\boldsymbol{v}$, i.e. in the negative direction of the x-axis. Let $l = CD$, and γ be the angle between the axes Cx and $C\xi$. Then the moment of the force $\boldsymbol{R}$ about C is equal to $Rl \sin\gamma$, and the elementary work of the force $\boldsymbol{R}$ will be

$$\delta A = Rl \sin\gamma \delta\gamma = -\delta\left(Rl \cos\gamma\right)$$

---

* R is a function of v, $R = f(v)$, and so from $v = \text{const}$ it follows that $R = \text{const}$.

therefore, for the force potential we can take the function*

$$\Pi = Rl\cos\gamma = Rl\cos\alpha\cos\beta$$

Oscillations of the axis of the shell are characterized by change in the angles α and β. To determine the stability of rotational motion of the shell we will proceed from three integrals of motion:

$$1)\ T+\Pi = \text{const};\quad 2)\ G_x = \text{const};\quad 3)\ G_\xi = Ap = \text{const}$$

The first integral is the energy integral; G_x and G_ξ are projections of the kinetic moment G_C on the axes Cx and $C\xi$. The constancy of G_x, when the system is in motion, follows from the fact that the moment of the force $\boldsymbol{R}$ about the Cx-axis is zero. The third integral expresses the constancy of the generalized momentum Ap, which corresponds to the cyclic coordinate φ. Note that (see Fig. 46)

$$G_x = G_\xi\cos(x\xi) + G_\eta\cos(x\eta) + G_\zeta\cos(x\zeta) = Ap\cos(x\xi) + Bq\cos(x\eta) +$$

$$+Br\cos(x\zeta) = Ap\cos\alpha\cos\beta + B(\dot\beta\sin\alpha - \dot\alpha\cos\alpha\cos\beta\sin\beta)^{**}$$

Combining the first two integrals with the third and using the expressions obtained for T, Π, and G_x, we find the following two integrals of motion W_1 and W_2 that vanish for $\alpha=\beta=\dot\alpha=\dot\beta=0$:

$$W_1 = \frac{1}{2}B(\dot\alpha^2\cos^2\beta + \dot\beta^2) + Rl(\cos\alpha\cos\beta - 1) = \text{const},$$

$$W_2 = B(\dot\beta\sin\alpha - \dot\alpha\cos\alpha\cos\beta\sin\beta) + Ap(\cos\alpha\cos\beta - 1) = \text{const}$$

We shall seek the fixed-sign linear combination of integrals of motion $W_1 - \lambda W_2$. First, determine in W_1 and W_2 the terms of lowest degree in the small quantities α, β, $\dot\alpha$, $\dot\beta$:

$$W_1 = \frac{1}{2}B(\dot\alpha^2+\dot\beta^2) - \frac{1}{2}Rl(\alpha^2+\beta^2) + \ldots$$

$$W_2 = B(\dot\beta\alpha - \dot\alpha\beta) - \frac{1}{2}Ap(\alpha^2+\beta^2) + \ldots$$

then

$$W_1 - \lambda W_2 = \frac{1}{2}[B\dot\alpha^2 + 2B\lambda\dot\alpha\beta + (Ap\lambda - Rl)\beta^2] +$$

$$+\frac{1}{2}[B\dot\beta^2 - 2B\lambda\dot\beta\alpha + (Ap\lambda - Rl)\alpha^2] + \ldots$$

So that each expression in the square brackets should be positive definite, it is sufficient that the following inequality be fulfilled:

$$B^2\lambda^2 < B(Ap\lambda - Rl)$$

---

* On a sphere with centre at C, the arcs α, β, γ form a rectangular spherical triangle with "sides" α, β and "hypotenuse" γ. For such a triangle, the formula $\cos\gamma = \cos\alpha\cos\beta$ holds. The truth of the formula follows from elementary geometrical reasoning.

** The arcs α, $\frac{\pi}{2}+\beta$ and $(x\zeta)$ form a rectangular spherical triangle with "sides" α and $\frac{\pi}{2}+\beta$, therefore

$$\cos(x\zeta) = \cos\alpha\cos\left(\frac{\pi}{2}+\beta\right) = -\cos\alpha\sin\beta$$

Cancelling out B and rearranging, we get

$$B\lambda^2 - Ap\lambda + Rl < 0$$

For the last inequality to hold for a certain real λ, it is necessary that the quadratic trinomial on the left-hand side of this inequality have real roots, i.e. there must be the inequality

$$A^2p^2 > 4BRl$$

This is the condition that ensures the stability of rotational motion of the shell (for "deviations" α, β, $\dot{\alpha}$, $\dot{\beta}$) because if this condition is fulfilled it is possible to select a real value of λ for which the integral $W_1 - \lambda W_2$ will have, at $\alpha = \beta = \dot{\alpha} = \dot{\beta} = 0$, a strict minimum equal to zero.

37. Stability of Linear Systems

In the preceding section it was shown that the investigation of the stability of any process defined by a system of ordinary differential equations* reduces to investigation of the stability of the zero solution of a system of deviation equations.

Let the differential deviation equations be linear and have constant coefficients

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} a_{ik}x_k \quad (i = 1, \ldots, n) \tag{1}$$

We shall seek a special solution of this system in the form

$$x_i = u_i e^{\lambda t} \ \left(i = 1, \ldots, n;\ \sum_{i=1}^{n} |u_i|^2 > 0\right) \tag{2}$$

Putting expressions (2) into equations (1) and cancelling out $e^{\lambda t}$, we get relations that connect the desired quantities u_i and λ:

$$\sum_{k=1}^{n} a_{ik}u_k = \lambda u_i \quad (i = 1, \ldots, n) \tag{3}$$

or

$$\sum_{k=1}^{n} (a_{ik} - \lambda\delta_{ik})\, u_k = 0 \quad (i = 1, \ldots, n) \tag{4}$$

where δ_{ik} is Kronecker's symbol

$$(\delta_{ik} = 1 \text{ at } i = k,\ \delta_{ik} = 0 \text{ at } i \neq k)$$

Since in the sought-for solution (2), at least one of the constants u_i must be different from zero, the determinant of the system of homogeneous equations (4) must be zero:

---

* It is known that such a system may ordinarily be written as a system of first-order differential equations solved for the derivatives. This notation is always possible for a system of Lagrange equations (see Sec. 7).

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0 \qquad (5)$$

Thus, for determining λ we obtained an algebraic equation in λ of degree n.

Equation (5) is called the *characteristic* or *secular* equation for the matrix of coefficients

$$A = \left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{matrix} \right\|$$

The roots of the characteristic equation are called *characteristic numbers* of the matrix A.

Taking for λ some characteristic number of the matrix A, we will find the constants u_i of the system of linear equations (4) that correspond to this number.

It will be convenient for what is to follow to introduce matrix notation both for the initial system (1) and for the systems of the relations (2) and (3).

We introduce the column vectors

$$x = \left\| \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{matrix} \right\|, \qquad u = \left\| \begin{matrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{matrix} \right\|$$

Then, in place of (1) to (3) and (5) we can write*

$$\frac{dx}{dt} = Ax, \qquad (1')$$

$$x = ue^{\lambda t}, \qquad (2')$$

$$Au = \lambda u \quad (u \neq 0), \qquad (3')$$

$$\det(A - \lambda E) = 0 \qquad (5')$$

where, $E = \| \delta_{ik} \|_{i,k=1}^{n}$ is the unit matrix.

---

[*] When multiplying the square matrix A by column x, we multiply elements of the ith row of A by the corresponding element of column x and add all these products. The sum thus obtained is the ith element of the column product Ax. The derivative of the column $\frac{dx}{dt}$ is obtained by differentiating each element of the column x.

The column $u \neq 0$ that, together with the number λ, satisfies the relation (3′) is called the *eigenvector* of the matrix A corresponding to the characteristic number λ. Thus, in each solution of the system (1′), having the form (2′), λ is the characteristic number of the matrix A, and u is the corresponding eigenvector.

We first consider the case when the characteristic equation (5) has n different roots λ_k $(k = 1, \ldots, n)$. To each characteristic number λ_k there corresponds an eigenvector u_k and a particular solution of the system (1) of the form $u_k e^{\lambda_k t}$. A linear combination of these solutions with arbitrary constant coefficients

$$x = \sum_{k=1}^{n} C_k u_k e^{\lambda_k t} \tag{6}$$

will again be a solution of the system (1).

To show that formula (6) embraces all solutions of (1), we shall first prove that the column vectors $u_1, u_2, \ldots, u_n$ which correspond to distinct characteristic numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ are linearly independent.

Let

$$\sum_{k=1}^{n} c_k u_k = 0 \tag{7}$$

Multiply both sides of (7) from the left by the matrix A. Then, using the equalities

$$A u_k = \lambda_k u_k \quad (u_k \neq 0,\ k = 1, \ldots, n)$$

we find

$$\sum_{k=1}^{n} \lambda_k c_k u_k = 0 \tag{8}$$

Eliminate the constant c_1 from (7) and (8):

$$\sum_{j=2}^{n} (\lambda_j - \lambda_1) c_j u_j = 0 \tag{9}$$

Again multiply this equality from the left by A and use the equality obtained, together with (9), to eliminate c_2 and so forth. We finally get

$$(\lambda_n - \lambda_1)(\lambda_n - \lambda_2) \ldots (\lambda_n - \lambda_{n-1}) c_n u_n = 0 \tag{10}$$

whence $c_n = 0$. Since all the terms in (7) are equivalent, it follows that

$$c_1 = c_2 = \ldots = c_n = 0$$

i.e. there is no relation of the form (7) between the eigenvectors $u_1, u_2, \ldots, u_n$, and these vectors are linearly independent.

Putting $t = 0$ in (6), we have B

$$x_0 = \sum_{k=1}^{n} C_k u_k \tag{11}$$

With an arbitrarily given initial vector x_0, we can uniquely determine C_k $(k = 1, \ldots, n)$ from (11) by virtue of the linear independence of the vectors $u_1, \ldots, u_n$. Thus, (6) embraces the solutions of the system (1) that satisfy any initial conditions $x(0) = x_0$, that is to say, it embraces all solutions of the system (1).

In courses on the theory of differential equations it is proved that in the case of multiple roots, formula (6) becomes somewhat involved. So-called *secular* terms can appear in the formula that contain a polynomial in t in place of the constant vector u_k: $u_k + u_k' t + \ldots$. In the general case, an arbitrary solution of the system of differential equations (1) is determined by the formula

$$x = \sum_{k=1}^{n} C_k (u_k + u_k' t + \ldots) e^{\lambda_k t} \tag{12}$$

Important corollaries follow at once from formulas (6) and (12).

1°. If all the characteristic numbers of the matrix A have negative real parts, i.e.

$$\max_{1 \leqslant k \leqslant n} \operatorname{Re} \lambda_k = -\alpha < 0$$

then $\lim_{t \to \infty} x(t) = 0$, and the zero solution of the system of differential equations (1) is asymptotically stable.*

Now let $\operatorname{Re} \lambda_k > 0$ at least for one k. Then (1) has a nonzero solution $x = C_k u_k e^{\lambda_k t}$ which tends to infinity as $t \to \infty$. At the same time, the initial value $x_0 = C_k u_k$ (for $t = 0$) may be arbitrarily small, since C_k is an arbitrary constant. In this case, the solution $x = 0$ is unstable. We thus have the proposition:

2°. If at least one characteristic number λ_k of the matrix A has a positive real part $(\operatorname{Re} \lambda_k > 0)$, then the zero solution of the system of differential equations (1) is unstable.**

Example. The position of equilibrium of a linear oscillator in a medium with a resistance proportional to the first power of the velocity will be asymptotically stable. Indeed (see Example 3 on page 167), the differential equation of motion $m\ddot{x} + 2f\dot{x} + cx = 0$ $(m, c, f > 0)$ may be written in the form of a system of two first-order differential equations if we put $x_1 = x$, $x_2 = \dot{x}$:

---

* The number $\alpha > 0$ is sometimes called the *degree of stability*.

** If all $\operatorname{Re} \lambda_k \leqslant 0$ $(k = 1, \ldots, n)$ and if in at least one of these relations there is an equal sign, then the solution $x = 0$ will be stable, provided that in (12) there are no secular terms in all summands where $\operatorname{Re}\lambda_k = 0$. Otherwise, he solution $x = 0$ will be unstable.

$$\frac{dx_1}{dt} = x_2, \quad \frac{dx_2}{dt} = -\frac{c}{m} x_1 - 2\frac{f}{m} x_2$$

The characteristic equation

$$\begin{vmatrix} -\lambda & 1 \\ -\frac{c}{m} & -2\frac{f}{m} - \lambda \end{vmatrix} \equiv \lambda^2 + 2\frac{f}{m}\lambda + \frac{c}{m} = 0$$

has roots with negative real part $-\frac{f}{m}$, which is what ensures the asymptotic stability of the equilibrium position.

38. Stability in Linear Approximation

In the system of differential equations (non-linear!)

$$\frac{dx_i}{dt} = X_i(x_1, \ldots, x_n, t) \quad (i = 1, \ldots, n) \tag{1}$$

expand the right sides in a series of powers of the deviations $x_1, \ldots, x_n$:

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} a_{ik} x_k + f_i \quad (i = 1, \ldots, n) \tag{1'}$$

where f_i is the sum of all terms of the expansion X_i beginning with second-order terms in $x_1, \ldots, x_n$ $(i = 1, \ldots, n)$.

In the stationary case, a_{ik} are constant coefficients, and the functions f_i depend on $x_1, \ldots, x_n$ and do not depend on t. In the periodic case, a_{ik} are periodic functions of t with period τ, and the nonlinear terms $f_i = f_i(x_1, \ldots, x_n, t)$ are also periodic in t with period τ.

If in (1′) we reject all nonlinear terms f_i, we get a linear system of differential equations that is called a linear approximation for the nonlinear system (1).

At the end of last century, it was established in the investigations of Poincaré and Lyapunov that both in the stationary and the periodic case, one can judge about the stability of a zero solution of a nonlinear system (1) from the linear approximation; namely, from the asymptotic stability of the zero solution of a linear approximation there follows the asymptotic stability of the zero solution of a nonlinear system.* This proposition finds broad applications since the investigation of linear systems is substantially simpler than that of nonlinear systems.

---

* It was assumed here that the right-hand sides $X_i(x_1, \ldots, x_n, t)$ are continuous functions. At the present time it is known what is to be understood by linear approximation for discontinuous right members of X_i, and an appropriate criterion has been established for stability based on a linear approximation both in the periodic case and in certain nonperiodic cases.

We will confine ourselves to a consideration of the stationary case and here we will write the system (1′) in the matrix form in order to prove our assertion:

$$\frac{dx}{dt}=Ax+f(x) \tag{1''}$$

Here, $A=\|a_{ik}\|_{i,k=1}^{n}$ is a square matrix with constant elements, and $f(x)$ is a column with elements $f_i(x_1, \ldots, x_n)$ $(i=1, \ldots, n)$. Since by assumption the zero solution of a linear approximation is asymptotically stable, it follows (see Sec. 37) that all the characteristic numbers $\lambda_1, \ldots, \lambda_n$ of the matrix A have negative real parts

$$\max_{1\leqslant k\leqslant n} \operatorname{Re}\lambda_k=-\alpha<0 \tag{2}$$

Let us agree to denote the "length" of the column vector z with components $z_1, \ldots, z_n$ by $|z|$:

$$|z|=\Big(\sum_{k=1}^{n}|z_k|^2\Big)^{\frac{1}{2}} \tag{3}$$

Since each element of the column $f(x)$ begins with terms of second degree, it follows that

$$|f(x)|\leqslant\varepsilon|x| \tag{4}$$

where the constant $\varepsilon>0$ may be chosen arbitrarily small if we confine the variation of the variables $x_1, \ldots, x_n$ to a sufficiently small neighbourhood $|x|<\Delta$.

The proof* of the assertion of Poincaré and Lyapunov may be built upon the following lemma:

Lemma. Using the linear nonsingular transformation of variables

$$x=Uy \quad (\det U\neq 0) \tag{5}$$

the linear system of differential equations $\frac{dx}{dt}=Ax$ may be reduced to the "triangular" form **

$$\left.\begin{aligned}
\frac{dy_1}{dt}&=\lambda_1y_1+b_{12}y_2+\ldots+b_{1n}y_n\\
\frac{dy_2}{dt}&=\qquad\quad\lambda_2y_2+\ldots+b_{2n}y_n\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\frac{dy_n}{dt}&=\qquad\qquad\qquad\qquad\lambda_ny_n
\end{aligned}\right\} \tag{6}$$

---

[*] See J. J. Stoker, *Nonlinear Vibrations in Mechanical and Electrical Systems*, London, Interscience Publishers, 1950.

[**] As a result of the transformation (5), the matrix of coefficients A is replaced by the matrix $U^{-1}AU$, which is of triangular form.

where $\lambda_1, \ldots, \lambda_n$ are characteristic numbers of the matrix A, while the moduli of the "nondiagonal" coefficients b_{ik} $(i < k)$ may be made arbitrarily small by a suitable choice of transformation (5). *

The transformation (5) is applicable to the nonlinear system (1″). The system (1″) will be written in new variables as follows:

$$\left.\begin{aligned} \frac{dy_1}{dt} &= \lambda_1 y_1 + b_{12} y_2 + \ldots + b_{1n} y_n + g_1(y) \\ \frac{dy_2}{dt} &= \qquad\quad \lambda_2 y_2 + \ldots + b_{2n} y_n + g_2(y) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{dy_n}{dt} &= \qquad\qquad\qquad\qquad \lambda_n y_n + g_n(y) \end{aligned}\right\} \tag{7}$$

where the column $g(y)$ with elements $g_1(y), \ldots, g_n(y)$ is defined by the equality

$$g(y) = U^{-1} f(Uy) \tag{8}$$

and satisfies the inequality**

$$|g(y)| < \eta |y| \tag{9}$$

where the number η (like the number ε) may be made arbitrarily small by choosing a sufficiently small neighbourhood $|x| < \Delta$ (and, respectively, $|y| < \Delta_1$).

Then, in accordance with equations (7) and inequalities (2) and (9), we find***

$$|y| \frac{d|y|}{dt} = \frac{1}{2} \frac{d}{dt} |y|^2 = \frac{1}{2} \frac{d}{dt} \sum_{i=1}^{n} y_i \bar{y}_i = \frac{1}{2} \sum_{i=1}^{n} \left(y_i \frac{d\bar{y}_i}{dt} + \bar{y}_i \frac{dy_i}{dt} \right) =$$

$$= \sum_{i=1}^{n} \operatorname{Re} \lambda_i |y_i|^2 + \frac{1}{2} \sum_{i>k} (b_{ik} y_k \bar{y}_i + \bar{b}_{ik} \bar{y}_k y_i) +$$

$$+ \frac{1}{2} \sum_{i=1}^{n} (g_i \bar{y}_i + \bar{g}_i y_i) \leqslant (-\alpha + \delta + \eta) |y|^2 \quad \left(\delta = \sum_{i<k} |b_{ik}| \right)$$

---

[*] The proof of this lemma is given on pages 196-197.

[**] If we define the norm of the matrix $A = \| a_{ik} \|_{i,\,k=1}^{n}$ by the equality $\| A \| = \left(\sum_{i,\,k=1}^{n} |a_{ik}|^2 \right)^{\frac{1}{2}}$ it will be easy to verify the truth of the inequality

$$|Ax| \leqslant \| A \| \, |x|$$

Therefore, from inequality (4) and equality (8) it follows that

$$|g(y)| \leqslant \| U^{-1} \| \, |f(Uy)| \leqslant \varepsilon \| U^{-1} \| \, \| U \| \, |y|$$

and, thus, in inequality (9) we can put

$$\eta = \| U \| \, \| U^{-1} \| \, \varepsilon.$$

[***] By $\bar{y}_k$ we denote a number that is complex conjugate to y_k $(k = 1, \ldots, n)$.

that is,

$$\frac{d|y|}{dt} \leqslant (-\alpha+\delta+\eta)|y|$$

whence

$$|y| \leqslant |y_0| e^{(-\alpha+\delta+\eta)(t-t_0)} \quad (y_0 = y(t_0)) \tag{10}$$

Choose positive numbers δ and η so that the inequalities $\alpha-\delta>0$ and $\eta<\alpha-\delta$ hold; then from (10) it follows that

$$|y| \leqslant |y_0| \text{ and } \lim_{t\to\infty} y(t)=0 \tag{11}$$

i.e. the solution $y = 0$ of the system (7) is asymptotically stable. But the vectors x and y are connected by the linear transformation (5); therefore the solution $x = 0$ of the system (1) is asymptotically stable.*

Proof of the lemma. First we shall show that with the aid of a transformation $x = U^{(1)}z$ of the form (5) it is possible to reduce the system of differential equations

$$\frac{dx}{dt} = Ax \tag{12}$$

to a form in which: (a) the first variable z_1 does not appear on the right of all equations beginning with the second, and (b) in the first equation the coefficient of z_1 is equal to the characteristic number λ_1 of the matrix A; i.e. to the following form:

$$\left.\begin{aligned}
\frac{dz_1}{dt} &= \lambda_1 z_1 + b'_{12}z_2 + \ldots + b'_{1n}z_n, \\
\frac{dz_2}{dt} &= \qquad\quad b'_{22}z_2 + \ldots + b'_{2n}z_n, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
\frac{dz_n}{dt} &= \qquad\quad b'_{n2}z_2 + \ldots + b'_{nn}z_n
\end{aligned}\right\} \tag{13}$$

For this purpose, it is sufficient, for the first column of the matrix $U^{(1)}$, to take the eigenvector u_1 which corresponds to the characteristic number λ_1 ($Au_1 = \lambda_1 u_1$, $u_1 \neq 0$) and choose the other columns of the matrix $u_2, \ldots, u_n$ so that together with u_1 they are linearly independent (then $\det U^{(1)} \neq 0$).

Indeed, the transformation $x = U^{(1)}z$ may be also written as follows:

$$x = u_1 z_1 + u_2 z_2 + \ldots + u_n z_n \tag{14}$$

where $z_1, \ldots, z_n$ are the coordinates of the vector x in the basis $u_1, u_2, \ldots, u_n$. The system (12) has the solution

$$x = u_1 e^{\lambda_1 t} \tag{15}$$

---

* In accordance with footnote ** on the preceding page, from the equality (5) follows the inequality

$$|x| \leqslant \|U\|\,|y| \leqslant \|U\|\,|y_0| \leqslant \|U\|\,\|U^{-1}\|\,|x_0|.$$

Therefore, the transformed system of differential equations

$$\frac{dz_i}{dt}=\sum_{k=1}^{n} b'_{ik}z_k \qquad (i=1, \ldots, n) \tag{16}$$

has the solution by (14),

$$z_1=e^{\lambda_1 t}, \quad z_2=\ldots=z_n=0 \tag{15'}$$

which is only possible when $b'_{11}=\lambda_1$ and $b'_{21}=\ldots=b'_{n1}=0$, i.e. when the system (16) has the form (13).

Since in a linear nonsingular transformation, the characteristic equation of the matrix A does not change,* the matrix $\| b_{jk} \|_2^n$ has as its characteristic numbers the remaining $n-1$ characteristic numbers $(\lambda_2, \ldots, \lambda_n)$ of the matrix A.

Applying a similar transformation to the system of the last $n-1$ equations (13) and so forth, we can finally, with the aid of a nonsingular linear transformation, reduce the initial system of differential equations to the form

$$\left.\begin{aligned} \frac{dz_1}{dt} &= \lambda_1 z_1 + b'_{12}z_2 + \ldots + b'_{1n}z_n, \\ \frac{dz_2}{dt} &= \qquad\quad \lambda_2 z_2 + \ldots + b'_{2n}z_n, \\ & \ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{dz_n}{dt} &= \qquad\qquad\qquad\qquad \lambda_n z_n \end{aligned}\right\} \tag{17}$$

Finally, we perform the transformation of variables

$$z_k=\mu^k y_k \quad (\mu>0;\ k=1, \ldots, n)$$

Then the system (17) will be replaced by the system (6) in which the off-diagonal coefficients $b_{ik}=\mu^{k-i}b'_{ik}$ $(i<k)$ may be made arbitrarily small in absolute value if the number $\mu>0$ is chosen sufficiently small. The lemma is proved.

39. Criteria of Asymptotic Stability of Linear Systems

In the preceding two sections it was established that in the stationary case, the zero solution of an arbitrary (nonlinear) system of differential equations in deviations is asymptotically stable if all the roots of the characteristic equation composed for the matrix of coefficients of a linear approximation have negative real parts.

---

* The characteristic equations of the matrices A and $U^{-1}AU$ (see footnote** on page 194) coincide since

$$U^{-1}(A-\lambda E)\,U=U^{-1}AU-\lambda E$$

and therefore

$$\det\,(U^{-1}AU-\lambda E)=\det U^{-1}\det\,(A-\lambda E)\det U=\det\,(A-\lambda E)$$

(E is the unit matrix).

Therefore, of great practical importance become the necessary and sufficient conditions that all the roots of the algebraic equation with real coefficients

$$f(\lambda) \equiv a_0\lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n = 0 \quad (a_0 > 0) \tag{1}$$

have negative real parts.

Denote by $\lambda_k\,(k = 1, \dots, g)$ the real, and by $r_j \pm is_j$ $\left(j = 1, \dots, \frac{n-g}{2}\right)$ the complex roots of the equation (1), and assume that in the complex plane all these roots lie to the left of the imaginary axis, i.e. that

$$\lambda_k < 0, \; r_j < 0 \;\left(k = 1, \dots, g;\; j = 1, \dots, \frac{n-g}{2}\right) \tag{2}$$

Then

$$f(\lambda) = a_0 \sum_{k=1}^{g} (\lambda - \lambda_k) \prod_{j=1}^{\frac{n-g}{2}} (\lambda - r_j - is_j)(\lambda - r_j + is_j) =$$

$$= a_0 \prod_{k=1}^{g} (\lambda - \lambda_k) \prod_{j=1}^{\frac{n-g}{2}} (\lambda^2 - 2r_j\lambda + r_j^2 + s_j^2) \tag{3}$$

Since, by the inequalities (2), each term in the last part of (3) has positive coefficients, it follows that in (1) all the coefficients are positive. The positivity of all coefficients is a necessary (for $a_0 > 0$) but by no means sufficient condition for all roots of the equation (1) to be located to the left of the imaginary axis.

In 1875 the English mechanician Routh, with whom the reader is already acquainted, produced an algorithm by means of which it is possible, using the coefficients of the polynomial $f(\lambda)$, to find out whether the polynomial is "stable", i.e. whether all its roots have negative real parts. In 1895, the German mathematician Hurwitz, independently of Routh, established the very same criterion in a modified form with the help of determinants (Hurwitz determinants):

$$\Delta_1 = a_1, \quad \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}, \quad \dots, \quad \Delta_n = \begin{vmatrix} a_1 & a_3 & a_5 & \dots \\ a_0 & a_2 & a_4 & \dots \\ 0 & a_1 & a_3 & \dots \\ 0 & a_0 & a_2 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & a \end{vmatrix} \tag{4}$$

(Here, $a_p = 0$ for $p > n$ everywhere.)

The Routh-Hurwitz condition. *For all roots of equation* (1) *to have negative real parts, it is necessary and sufficient that the following inequalities hold*:

$$\Delta_1 > 0,\ \Delta_2 > 0,\ \ldots,\ \Delta_n > 0 \tag{5}$$

If the coefficients of (1) are given as numbers, then the conditions (5) are readily verified. But if the coefficients of (1) contain literal parameters, then the computation of the determinants Δ_k for large k is already complicated.

It is therefore of interest to examine other conditions established in 1914 by the French mathematicians Liénard and Chipart. In these conditions, the number of determinantal inequalities is roughly half that in the conditions (5) of Routh and Hurwitz.

The Liénard-Chipart conditions. For a polynomial $f(\lambda) \equiv a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n$, when $a_0 > 0$, to have all roots with negative real parts, it is necessary and sufficient that

(1) all coefficients of the polynomial $f(\lambda)$ be positive

$$a_1 > 0,\ a_2 > 0,\ \ldots,\ a_n > 0 \tag{6}$$

(2) the determinantal inequalities

$$\Delta_{n-1} > 0,\ \Delta_{n-3} > 0,\ \ldots \tag{7}$$

be valid. (Here, as before, Δ_k denotes the Hurwitz determinant of the kth order.)

We shall now get acquainted with the geometric criterion of stability.

In the equality*

$$f(\lambda) = a_0 \prod_{k=1}^{n} (\lambda - \lambda_k)$$

replace λ by $i\omega$ and let ω vary from $-\infty$ to $+\infty$. Compute the appropriate increment of the angle $\theta = \arg f(i\omega)$:

$$\Delta_{-\infty}^{+\infty}\theta(\omega) = \sum_{k=1}^{n} \Delta_{-\infty}^{+\infty} \arg(i\omega - \lambda_k)$$

Now note that** (Fig. 47)

$$\Delta_{-\infty}^{+\infty} \arg(i\omega - \lambda_k) = \begin{cases} \pi \text{ if } \operatorname{Re}\lambda_k < 0 \\ -\pi \text{ if } \operatorname{Re}\lambda_k > 0 \end{cases}$$

---

* Here, the n roots of the polynomial $f(\lambda)$ are denoted by $\lambda_1, \ldots, \lambda_n$.

** We assume here that not a single one of the roots λ_k lies on the imaginary axis.

Therefore, denoting by l and r the number of roots lying to the left and, respectively, to the right of the imaginary axis ($l + r = n$), we will have

$$\Delta_{-\infty}^{+\infty}\theta(\omega) = (l - r)\pi$$

We consider the curve described by the affix* of the complex number $f(i\omega)$ as ω varies from $-\infty$ to $+\infty$. This curve decomposes into two branches: on one $\omega > 0$, on the other $\omega < 0$. One branch is obtained from the other by a mirror mapping about the real axis, since $f(i\omega)$ and $f(-i\omega)$ are complex conjugate numbers. For this reason, denoting by Δ_0^∞ the increment as ω varies from 0 to ∞, we obtain

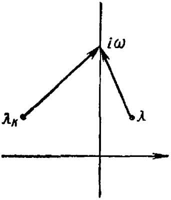

Fig. 47

$$\Delta_0^\infty\theta(\omega) = \frac{1}{2}\Delta_{-\infty}^{+\infty}\theta(\omega) = (l - r)\frac{\pi}{2} \tag{8}$$

Whence it is seen that all the roots will be located to the left of the imaginary axis ($l = n$, $r = 0$) when, and only when,

$$\Delta_0^\infty\theta(\omega) = n\frac{\pi}{2} \tag{9}$$

Geometric criterion of stability.** For a polynomial $f(\lambda)$ to be stable, i.e. for all its roots to be located to the left of the imaginary axis, it is necessary and sufficient:

(1) that the hodograph of $f(i\omega)$, as ω varies from 0 to ∞, should not pass through the zero point*** and

(2) that for this hodograph

$$\Delta_0^\infty\Theta = n\frac{\pi}{2}$$

---

* The affix of a complex number z is an appropriate point in the complex z-plane.

** This criterion was first applied by A. V. Mikhailov for investigating automatic control systems. For this reason, in the technical literature, the geometric criterion of stability is frequently called Mikhailov's criterion.

*** Condition (1) signifies that $f(\lambda)$ does not have any purely imaginary roots.

where n is the degree of the polynomial $f(\lambda)$ (see Fig. 48 for $n = 6$).

Note that for a stable polynomial* the argument θ varies monotonically as ω varies from 0 to ∞. This follows from the formula

$$\theta(\omega) = \sum_{k=1}^{n} \arg(i\omega - \lambda_k)$$

since for such a polynomial each summand on the right-hand side is a monotonic increasing function of ω.

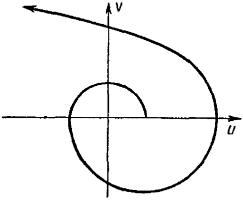

Fig. 48

Example. Let $f(\lambda) = \lambda^5 + 5\lambda^4 + 10\lambda^3 + 11\lambda^2 + 7\lambda + 2$. Then $f(i\omega) = U(\omega) + iV(\omega)$ where

$$U(\omega) = 5\omega^4 - 11\omega^2 + 2, \quad V(\omega) = \omega(\omega^4 - 10\omega^2 + 7)$$

To construct the hodograph of $f(i\omega)$ we note that $U(0) = 2$ and $V(\omega)$ vanishes at $\omega = 0$ and at $\omega = \omega_1$, $\omega = \omega_2$ $(0 < \omega_1 < \omega_2)$; the squares ω_1^2 and ω_2^2 are determined from the quadratic equation

$$\omega_1^2 = 5 - \sqrt{18} \approx 0.76; \quad \omega_2^2 = 5 + \sqrt{18} \approx 9.24$$

It will readily be seen that $U(\omega_1) < 0$, $U(\omega_2) > 0$. Besides, $V'(0) = 7 > 0$.

Thus, for $\omega = 0$ the hodograph begins on the positive real axis, goes upwards at first, intersects the positive imaginary axis, then the negative real axis (at $\omega = \omega_1$), the negative imaginary axis, and, finally, again the positive real axis (at $\omega = \omega_2$). When $\omega > \omega_2$, the hodograph does not intersect the coordinate axes and goes off to infinity in the fifth quadrant $(U > 0,\ V > 0)$, since $n = 5$. Here, $\tan\theta = \frac{V(\omega)}{U(\omega)} \to +\infty$ when $\omega \to +\infty$. Hence,

$$\Delta_0^{\infty}\theta = 5\frac{\pi}{2}$$

i.e. $f(\lambda)$ is a stable polynomial.

We could arrive at the same conclusion by proceeding from the Liénard-Chipart criterion, since all the coefficients in $f(\lambda)$ are positive and

$$\Delta_2 = \begin{vmatrix} 5 & 11 \\ 1 & 10 \end{vmatrix} > 0, \quad \Delta_4 = \begin{vmatrix} 5 & 11 & 2 & 0 \\ 1 & 10 & 7 & 0 \\ 0 & 5 & 11 & 2 \\ 0 & 1 & 10 & 7 \end{vmatrix} > 0$$

---

* The stable polynomial is also called the *Hurwitz polynomial.*

CHAPTER 6

Small Oscillations

40. Small Oscillations of a Conservative System

If at an initial instant of time the position of a scleronomic system is chosen sufficiently close to a position of stable equilibrium and the initial velocities are sufficiently small in absolute value, then throughout all the motion both the deviations from the position of equilibrium and the generalized velocities will be small in absolute value. This circumstance permits retaining, in the differential equations of motion, only the linear terms in the deviations and the velocities, and higher-order terms may be rejected. Then the differential equations of motion become linear, i.e. the problem is "linearized". In this section we will consider the linearization of equations of motion for a conservative system.

The kinetic and the potential energy of a conservative system with n degrees of freedom are expressed in terms of the independent coordinates q_i and the generalized velocities $\dot{q}_i$ ($i = 1, \ldots, n$) in the following manner:

$$T = \frac{1}{2} \sum_{i,k=1}^{n} a_{ik}(q_1, \ldots, q_n)\, \dot{q}_i \dot{q}_k, \quad \Pi = \Pi(q_1, \ldots, q_n) \tag{1}$$

As in the preceding chapter, we assume that the coordinate origin $q_1 = \ldots = q_n = 0$ is a position of equilibrium and that in this position $\Pi = 0$. Expand the coefficients $a_{ik}(q_1, \ldots, q_n)$ in series in powers of the coordinates:

$$a_{ik}(q_1, \ldots, q_n) = a_{ik} + \ldots \quad (i, k = 1, \ldots, n) \tag{2}$$

where $a_{ik} = a_{ik}(0, \ldots, 0)$. ($a_{ik} = a_{ki}$; i, $k = 1, \ldots, n$) are constants. Putting these expressions for the coefficients into formula (1) for the kinetic energy, we have

$$T = \frac{1}{2} \sum_{i,k=1}^{n} a_{ik} \dot{q}_i \dot{q}_k + (**) \tag{3}$$

where by (**) we denote the sum of the terms of third and higher orders in q_i and $\dot{q}_i$ ($i = 1, \ldots, n$).

Also expand the potential energy in a power series of the coordinates:

$$\Pi = \Pi_0 + \sum_{i=1}^{n} \left(\frac{\partial \Pi}{\partial q_i}\right)_0 q_i + \frac{1}{2} \sum_{i,\,k=1}^{n} \left(\frac{\partial^2 \Pi}{\partial q_i\, \partial q_k}\right)_0 q_i q_k + (**)$$

By hypothesis, $\Pi_0 = 0$. Besides, the generalized forces in the equilibrium position are zero:

$$Q_i^0 = \left(\frac{\partial \Pi}{\partial q_i}\right)_0 = 0 \quad (i = 1, \ldots, n)$$

Therefore, introducing the notation

$$c_{ik} = \left(\frac{\partial^2 \Pi}{\partial q_i\, \partial q_k}\right)_0 \quad (c_{ik} = c_{ki};\ i,\ k = 1, \ldots, n) \tag{4}$$

we can also represent the potential energy in the form

$$\Pi = \frac{1}{2} \sum_{i,\,k=1}^{n} c_{ik} q_i q_k + (**) \tag{5}$$

Dropping terms of the third and higher orders of smallness in $\dot{q}_i$ and q_k in formulas (3) and (5), we can represent the kinetic and potential energies as quadratic forms with constant coefficients:

$$T = \frac{1}{2} \sum_{i,\,k=1}^{n} a_{ik} \dot{q}_i \dot{q}_k \qquad \Pi = \frac{1}{2} \sum_{i,\,k=1}^{n} c_{ik} q_i q_k \tag{6}$$

where $a_{ik} = a_{ki}$, $c_{ik} = c_{ki}$ $(i,\ k = 1, \ldots, n)$.

From the physical meaning of the kinetic energy it is clear that always $T \geqslant 0$. Since we assume that the position of equilibrium is not a singular point*, then always $T > 0$ if only not all generalized velocities are simultaneously equal to zero, i.e. the quadratic form $\sum_{i,\,k=1}^{n} a_{ik} \dot{q}_i \dot{q}_k = 2T$ is positive definite

$$\sum_{i,\,k=1}^{n} a_{ik} \dot{q}_i \dot{q}_k > 0 \quad \left(\sum_{i=1}^{n} \dot{q}_i^2 > 0\right) \tag{7}$$

Further, in order to ensure stability of the given position of equilibrium, we demand that (in accord with the Lagrange theorem) the potential energy have a strict minimum in the equilibrium position. Since $\Pi_0 = 0$, this means that in a certain neighbourhood of the origin

$$\sum_{i,\,k=1}^{n} c_{ik} q_i q_k > 0 \quad \left(\sum_{i=1}^{n} q_i^2 > 0\right) \tag{8}$$

---

* See footnote on page 47.

But the quadratic form (8) is a quadratic homogeneous function in the coordinates. Hence, inequality (8) holds throughout the space, with the exception of the origin where this form vanishes. In other words, the potential energy is now represented as a positive definite quadratic form in the coordinates.*

Form the Lagrange equations by proceeding from the expressions (6) for T and Π:

$$\sum_{k=1}^{n} (a_{ik}\ddot{q}_k + c_{ik}q_k) = 0 \quad (i = 1, \ldots, n) \tag{9}$$

We shall seek a particular solution of this system of linear differential equations in the form

$$q_i = u_i \sin(\omega t + \alpha) \quad (i = 1, \ldots, n) \tag{10}$$

i.e. in the form of harmonic oscillations with one and the same frequency ω and the same constant α for all coordinates.

Substituting expressions (10) for q_i in the differential equations (9) and setting

$$\lambda = \omega^2 \tag{11}$$

we obtain the following system of algebraic equations [after cancelling out $\sin(\omega t + \alpha)$] that are linear in the amplitudes u_i:**

$$\sum_{k=1}^{n} (c_{ik} - \lambda a_{ik}) u_k = 0 \quad (i = 1, \ldots, n) \tag{12}$$

Since all the amplitudes u_i of the sought-for oscillation must not vanish, the determinant of the system of homogeneous equations (12) must be equal to zero:

$$\begin{vmatrix} c_{11} - \lambda a_{11} & c_{12} - \lambda a_{12} & \ldots & c_{1n} - \lambda a_{1n} \\ c_{21} - \lambda a_{21} & c_{22} - \lambda a_{22} & \ldots & c_{2n} - \lambda a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ c_{n1} - \lambda a_{n1} & c_{n2} - \lambda a_{n2} & \ldots & c_{nn} - \lambda a_{nn} \end{vmatrix} = 0 \tag{13}$$

---

* Of course, cases are possible when the function $\Pi(q_1, \ldots, q_n)$ prior to rejection of terms (**) has a strict minimum at the origin, and after rejection of these terms has a nonstrict minimum at the same point (for example, $\Pi = c^2 (q_1 + \ldots + q_n)^2 + d^2 (q_1^4 + \ldots + q_n^4)$, $c > 0$, $d > 0$. But we shall consider such cases as special and exclude them. In these special cases, rejection of the terms (**) in the expression for Π is not justified, since it can radically distort the picture of the motion.

** We call u_i the amplitude of harmonic oscillation (10) of the coordinate q_i, though actually the absolute value $|u_i|$ is the amplitude; the initial phase (at $t = 0$) of harmonic oscillations (10) is either α (at $u_i > 0$) or $-\alpha$ (at $u_i < 0$).

Expanding the determinant, we get on the left-hand side a polynomial of degree n in λ. Thus, the square of the frequency $\lambda = \omega^2$ of the desired harmonic solution (10) must satisfy the nth degree algebraic equation (13), which is called the *secular equation* or *frequency equation.*

To every root λ of (13) there corresponds a particular solution (10) (for an arbitrary constant α) of the system of differential equations (9). In this solution, $\omega = \sqrt{\lambda}$.

Write the above formulas in matrix form. We introduce two symmetric positive definite matrices*

$$A = \| a_{ik} \| = \left\| \begin{matrix} a_{11} & \dots & a_{1n} \\ . & . & . \\ a_{n1} & \dots & a_{nn} \end{matrix} \right\|, \quad C = \| c_{ik} \| = \left\| \begin{matrix} c_{11} & \dots & c_{1n} \\ . & . & . \\ c_{n1} & \dots & c_{nn} \end{matrix} \right\| \tag{14}$$

and the column vectors

$$\boldsymbol{q} = \left\| \begin{matrix} q_1 \\ . \\ \vdots \\ . \\ q_n \end{matrix} \right\|, \quad \boldsymbol{u} = \left\| \begin{matrix} u_1 \\ . \\ . \\ . \\ u_n \end{matrix} \right\| \tag{15}$$

($\boldsymbol{u}$ is the amplitude vector). Then the system of differential equations (9) will be written as

$$A\ddot{\boldsymbol{q}} + C\boldsymbol{q} = 0 \tag{16}$$

The particular solution (10) will look like

$$\boldsymbol{q} = \boldsymbol{u} \sin(\omega t + \alpha) \tag{17}$$

and the result of substituting the solution (17) into equation (16), i.e. the system of algebraic equations (12), has the form

$$(C - \lambda A)\,\boldsymbol{u} = 0 \quad (\lambda = \omega^2) \tag{18}$$

The frequency equation will be written as

$$\det(C - \lambda A) = 0 \quad (\lambda = \omega^2) \tag{19}$$

So as to figure out whether the roots λ of the secular equation (19) are always real and positive, first consider some properties of quadratic forms with real coefficients.

---

* The symmetric matrix $A = \| a_{ik} \|_1^n$ is called positive definite if the quadratic form $\sum\limits_{i,\,k=1}^{n} a_{ik} q_i q_k$ corresponding to it is positive definite.

To each quadratic form $\sum_{i,k=1}^{n} a_{ik}u_iu_k$ there corresponds a certain bilinear form $\sum_{i,k=1}^{n} a_{ik}u_iv_k$ for which we introduce the abbreviated notation

$$A(u, v)=\sum_{i,k=1}^{n} a_{ik}u_iv_k$$

Then the quadratic form will be written as

$$A(u, u)=\sum_{i,k=1}^{n} a_{ik}u_iu_k$$

The following properties of a bilinear form are readily verifiable:

1°. $A(u_1+u_2, v)=A(u_1, v)+A(u_2, v)$.

2°. $A(\lambda u, v)=\lambda A(u, v)$ (λ is a scalar).

3°. $A(u, v)=A(v, u)$.*

We shall also show that for any complex vector u:

4°. $A(u, \bar{u})$ is a real number.**

Indeed, setting $u=v+iw$ (v and w are real column vectors) we find, by virtue of 1° to 3°:

$$A(u, \bar{u})=A(v+iw, v-iw)=A(v, v)-iA(v, w)+iA(w, v)+\\+A(w, w)=A(v, v)+A(w, w) \tag{20}$$

The last expression is obviously real.

From (20) it also follows that:

5°. If $A(u, u)$ is a positive definite quadratic form and $u\neq 0$ is an arbitrary complex vector, then

$$A(u, \bar{u})>0 \quad (u\neq 0) \tag{21}$$

Indeed, assuming $u=v+iw$ we have $A(v, v)\geqslant 0$, and $A(w, w)\geqslant 0$. In one of these relations we have the sign $>$, since from $u\neq 0$ follows at least one of the inequalities $v\neq 0$, $w\neq 0$. Then from equality (20) follows the equality (21).

Let us now prove:

6°. If λ is a root of the secular equation $\det(C-\lambda A)=0$, and u is the amplitude vector corresponding to it [see (18)]

$$Cu=\lambda Au \qquad (u\neq 0) \tag{22}$$

---

* Unlike equalities 1° and 2°, equality 3° holds only for a bilinear form with a symmetric matrix of coefficients.

** The bar marks the transition to complex conjugate quantities. Property 4° is valid only for a symmetric matrix with real elements.

then for any vector v

$$C(\boldsymbol{u}, \boldsymbol{v}) = \lambda A(\boldsymbol{u}, \boldsymbol{v}) \tag{23}$$

Indeed, in scalar notation, (22) takes the form

$$\sum_{k=1}^{n} c_{ik}u_k = \lambda \sum_{k=1}^{n} a_{ik}u_k \quad (i = 1, \ldots, n) \tag{22'}$$

Multiplying both sides of the ith equation of (22') by v_i and summing over i, we get

$$\sum_{i,\,k=1}^{n} c_{ik}v_i u_k = \lambda \sum_{i,\,k=1}^{n} a_{ik}v_i u_k$$

which is the equality (23).

We shall now show that *for any two amplitude vectors* $\boldsymbol{u}$ *and* $\boldsymbol{u}'$ *corresponding to different roots* λ *and* λ' ($\lambda \neq \lambda'$) *of the secular equation the following relation** *is valid*:

$$A(\boldsymbol{u}, \boldsymbol{u}') = 0 \tag{24}$$

Indeed, according to 6°, two equations hold:

$$C(\boldsymbol{u}, \boldsymbol{u}') = \lambda A(\boldsymbol{u}, \boldsymbol{u}'), \; C(\boldsymbol{u}, \boldsymbol{u}') = \lambda' A(\boldsymbol{u}, \boldsymbol{u}') \tag{25}$$

But $\lambda \neq \lambda'$. Therefore, from (25) follows relation (24).**

We now prove that *from the symmetrical nature of the matrices* A *and* C *and from the positive definiteness of the matrix* A *it follows that the secular equation* (13) [or (19)] *has only real roots.*

Indeed, let λ be a complex root of the secular equation ($\lambda \neq \bar{\lambda}$) and let a complex vector $\boldsymbol{u} \neq 0$ correspond to it. Then $\bar{\lambda}$ is also a root of the secular equation with amplitude vector $\bar{\boldsymbol{u}}$. Since $\lambda \neq \bar{\lambda}$, it follows from what has been proved that $A(\boldsymbol{u}, \bar{\boldsymbol{u}}) = 0$, but this contradicts the inequality (21).

If λ is real, then also the amplitude vector $\boldsymbol{u} \neq 0$ that corresponds to it may be chosen real. Now, putting $\boldsymbol{v} = \boldsymbol{u}$ in (23) and noting that

---

* If we introduce into n-dimensional space an A metric, i.e. if we take the square of the length of the vector $\boldsymbol{u}$ to be the magnitude of the quadratic form

$$A(\boldsymbol{u}, \boldsymbol{u}') = \sum_{i,\,k=1}^{n} a_{ik}u_i u_k$$

then $A(\boldsymbol{u}, \boldsymbol{u}')$ is the "scalar product" of the vectors $\boldsymbol{u}$ and $\boldsymbol{u}'$ in this metric. Therefore, (3) expresses the following property of amplitude vectors: amplitude vectors corresponding to different roots of the secular equation are always mutually orthogonal in the A metric.

** By virtue of (25), $A(\boldsymbol{u}, \boldsymbol{u}') = 0$ takes place along with $C(\boldsymbol{u}, \boldsymbol{u}') = 0$.

$A(\boldsymbol{u}, \boldsymbol{u}) > 0$, we find

$$\lambda = \frac{C(\boldsymbol{u}, \boldsymbol{u})}{A(\boldsymbol{u}, \boldsymbol{u})} \tag{26}$$

But in our case the quadratic form $C(\boldsymbol{u}, \boldsymbol{u}) = \sum\limits_{i,k=1}^{n} c_{ik}u_iu_k$ is also positive definite. Then not only $A(\boldsymbol{u}, \boldsymbol{u}) > 0$, but also $C(\boldsymbol{u}, \boldsymbol{u}) > 0$. Hence, $\lambda > 0$.

Thus, the secular equation (13) has n positive roots λ_j, to which there correspond real positive frequencies $\omega_j = \sqrt{\lambda_j}$ and real amplitude vectors $\boldsymbol{u}_j$ $(j = 1, \ldots, n)$.

Let us first consider the case when all the roots of the secular equation are distinct. To each λ_j there corresponds a particular solution

$$\boldsymbol{q} = \boldsymbol{u}_j \sin(\omega_j t + \alpha_j) \quad (\omega_j = \sqrt{\lambda_j}) \tag{27}$$

with an amplitude vector $\boldsymbol{u}_j$, the coordinates of which $u_{1j}, \ldots, u_{nj}$ must satisfy the system of linear equations

$$\sum_{k=1}^{n} (c_{ik} - \lambda a_{ik}) u_{kj} = 0 \quad (i = 1, \ldots, n) \tag{28}$$

or in matrix notation

$$(\boldsymbol{C} - \lambda_j \boldsymbol{A}) \boldsymbol{u}_j = 0 \tag{29}$$

Since the system of differential equations (9) [or (16)] is linear, the linear combination with constant coefficients of the solutions (27) is again a solution of this system. For this reason

$$\boldsymbol{q} = \sum_{j=1}^{n} C_j \boldsymbol{u}_j \sin(\omega_j t + \alpha_j) \quad (\omega_j = \sqrt{\lambda_j}; \quad j = 1, \ldots, n) \tag{30}$$

with arbitrary constants C_j, α_j $(j = 1, \ldots, n)$ is a solution of the system (9) or (16). We shall show that the formula (30) embraces all motions of the system.

We shall first prove that n amplitude vectors $\boldsymbol{u}_j$ $(j = 1, \ldots, n)$ are linearly independent.* Indeed, let

$$\sum_{j=1}^{n} c_j \boldsymbol{u}_j = 0.$$

Then for any fixed k $(1 \leqslant k \leqslant n)$

$$0 = A\left(\boldsymbol{u}_k, \sum_{j=1}^{n} c_j \boldsymbol{u}_j\right) = \sum_{j=1}^{n} c_j A(\boldsymbol{u}_k, \boldsymbol{u}_j) \tag{31}$$

---

* This was proved in Sec. 37 for the special case when $\boldsymbol{A}$ is a unit matrix.

But $A(u_k, u_j) = 0$ for $k \neq j$ and $A(u_k, u_j) > 0$ for $k = j$. Therefore, from equalities (31) it follows that

$$c_k = 0 \quad (k = 1, \ldots, n)$$

i.e. there can be no linear relationship between the vectors $u_1, \ldots, u_n$.

Now in formula (30) we choose the values of the arbitrary constants C_j, α_j in such manner that the following arbitrary preassigned initial conditions should hold:

$$q_i(0) = q_{i0}, \quad \dot{q}_i(0) = \dot{q}_{i0} \quad (i = 1, \ldots, n) \tag{32}$$

or in matrix notation

$$q(0) = q_0, \quad \dot{q}(0) = \dot{q}_0 \tag{33}$$

From formula (30) we get

$$\left.\begin{aligned} q_0 &= \sum_{j=1}^{n} C_j \sin \alpha_j u_j, \\ \dot{q}_0 &= \sum_{j=1}^{n} \omega_j C_j \cos \alpha_j u_j \end{aligned}\right\} \tag{34}$$

By virtue of the linear independence of the vectors u_j $(j = 1, \ldots, n)$, we determine uniquely from this the products $C_j \sin \alpha_j$ and $\omega_j C_j \cos \alpha_j$ and, consequently, since $\omega_j \neq 0$, we uniquely determine the values of the arbitrary constants C_j and α_j $(j = 1, \ldots, n)$.*

Thus, in the absence of multiple roots in the secular equation the formula (30) embraces all oscillations of the system.**

Now if the frequency equation has multiple roots, then it may be asserted that there will be at least m solutions of the form $u \sin(\omega t + \alpha)$, where m is the number of distinct roots λ_j of the secular equation.

Lagrange considered that in the case of multiple frequencies, the general solution of the system (9) is no longer representable in the form (30) and that on the right-hand side of (30) so-called secular terms of the form

$$(u + u't + u''t^2 + \ldots) \sin(\omega t + \alpha)$$

appear.

---

* The α_j are determined up to an additive a multiple of 2π.

** The following relations hold for the amplitude vectors u_j:

$$A(u_j, u_h) = \sum_{i,k=1}^{n} a_{ik} u_{ij} u_{kh} = 0 \quad (j \neq h;\ j, h = 1, \ldots, n)$$

However, Lagrange was in error. As Weierstrass demonstrated later, to each root λ of multiplicity p there corresponds exactly p linearly independent solutions of the system of linear equations (12), i.e. for each root λ_j of multiplicity p it is possible to find p linearly independent amplitude vectors. Thus, also in the case of multiple frequencies there exist n linearly independent amplitude vectors, and the formula (30) which is formed with their help yields a general solution in this case as well.

The oscillations

$$\boldsymbol{q} = C_j \boldsymbol{u}_j \sin(\omega_j t + \alpha_j) \quad (j = 1, \ldots, n) \tag{35}$$

that make up an arbitrary oscillation of a system are called *principal oscillations* of the system.

A rigorous derivation of the formula (30) for the general case (i.e. given multiple frequencies as well) with the aid of the so-called "normal" coordinates will be given in the next section. In this derivation, the case of multiple roots of the secular equation is not specifically isolated.

Example. Coupled pendulums. The points of suspension of two identical simple pendulums each of mass m and length l are located on one horizontal straight line. Points of these pendulums, at a distance h from the points of suspension ($0 < h \leqslant l$) are connected by a spring of rigidity γ; the spring is relaxed when the pendulums occupy the vertical position. It is required to determine the oscillation of the system in the vertical plane.

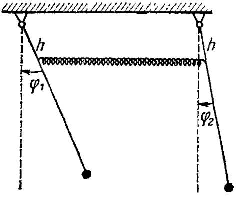

Fig. 49

For the independent coordinates we take the angles φ_1 and φ_2 formed by the pendulums and the vertical (Fig. 49). In the equilibrium position, $\varphi_1 = \varphi_2 = 0$. Up to higher-order infinitesimals, the extension of the spring is equal to $h \mid \sin \varphi_2 - \sin \varphi_1 \mid \cong h \mid \varphi_2 - \varphi_1 \mid$. Therefore, in the given case

$$T = \frac{1}{2} ml^2 (\dot{\varphi}_1^2 + \dot{\varphi}_2^2),$$

$$\Pi = mgl(1 - \cos\varphi_1) + mgl(1 - \cos\varphi_2) + \frac{\gamma h^2 (\varphi_2 - \varphi_1)^2}{2} + \ldots = \frac{1}{2} mgl(\varphi_1^2 + \varphi_2^2) + \frac{1}{2}\gamma h^2 (\varphi_2 - \varphi_1)^2 + \ldots$$

Retaining in Π only the quadratic terms, we finally get

$$T=\frac{1}{2}a\,(\dot{\varphi}_1^2+\dot{\varphi}_2^2)$$

$$\Pi=\frac{1}{2}c\,(\varphi_1^2+\varphi_2^2)-b\varphi_1\varphi_2$$

where

$$a=ml^2,\quad c=mgl+\gamma h^2$$

$$b=\gamma h^2$$

Write the equation of frequencies

$$\begin{vmatrix} c-\lambda a & -b \\ -b & c-\lambda a \end{vmatrix}=0 \qquad (\lambda=\omega^2)$$

and one of two equations for determining the amplitudes in the principal oscillations (these two equations are dependent):

$$(c-\lambda a)\,u_1-bu_2=0, \text{ i.e. } \frac{u_1}{u_2}=\frac{b}{c-\lambda a}$$

From the frequency equation we get

$$\omega_1^2=\lambda_1=\frac{c-b}{a}=\frac{g}{l}, \quad \omega_2^2=\lambda_2=\frac{c+b}{a}=\frac{g}{l}+2\frac{\gamma}{m}\frac{h^2}{l^2}$$

Accordingly, for the first principal oscillation $u_1=u_2=C_1$ and

$$\varphi_1=C_1\sin(\omega_1 t+\alpha_1),\quad \varphi_2=C_1\sin(\omega_1 t+\alpha_1) \quad (\varphi_1=\varphi_2)$$

for the second $u_1=-u_2=C_2$ and

$$\varphi_1=C_2\sin(\omega_2 t+\alpha_2),\quad \varphi_2=-C_2\sin(\omega_2 t+\alpha_2) \quad (\varphi_1=-\varphi_2)$$

In the first principal oscillation both pendulums are all the time in one phase, the spring is relaxed and the pendulums do not exert any effect on one another. In the second principal oscillation the pendulums are all the time in opposite phases.

Arbitrary oscillation is obtained by superimposition of the two principal oscillations:

$$\varphi_1=C_1\sin(\omega_1 t+\alpha_1)+C_2\sin(\omega_2 t+\alpha_2)$$

$$\varphi_2=C_1\sin(\omega_1 t+\alpha_1)-C_2\sin(\omega_2 t+\alpha_2)$$

41. Normal Coordinates

Two quadratic forms

$$\left.\begin{aligned} &A(q,q)=\sum_{i,k=1}^{n}a_{ik}q_iq_k \\ &\text{and} \\ &C(q,q)=\sum_{i,k=1}^{n}c_{ik}q_iq_k \end{aligned}\right\} \tag{1}$$

of which at least one, say $A(q,q)$, is positive definite, can always, by means of one and the same (nonsingular) transformation of variables,

$$q_i = \sum_{j=1}^{n} u_{ij}\theta_j \quad (i = 1, \ldots, n;\ \det (u_{ij})_{i,\,j=1}^{n} \neq 0) \tag{2}$$

be reduced to the "sum of squares"

$$A(\boldsymbol{q}, \boldsymbol{q}) = \sum_{j=1}^{n} \theta_j^2,\ C(\boldsymbol{q}, \boldsymbol{q}) = \sum_{j=1}^{n} \lambda_j\theta_j^2 \tag{3}$$

Here, all $\lambda_j > 0$, since (see Sec. 40) the form $C(\boldsymbol{q}, \boldsymbol{q})$ is also positive definite.

Since the generalized velocities $\dot{q}_i$ and $\dot{\theta}_j$ are connected by the same relations which connect the q_i and θ_j:

$$\dot{q}_i = \sum_{j=1}^{n} u_{ij}\dot{\theta}_j \quad (i = 1, \ldots, n),$$

it follows that in the first of the equalities (3) we can replace q_i and θ_j by $\dot{q}_i$ and $\dot{\theta}_j$, after which we get the following expressions for the kinetic energy and the potential energy:

$$\left.\begin{aligned} T &= \frac{1}{2} \sum_{i,\,k=1}^{n} a_{ik}\dot{q}_i\dot{q}_k = \frac{1}{2}\sum_{j=1}^{n} \dot{\theta}_j^2, \\ \Pi &= \frac{1}{2} \sum_{i,\,k=1}^{n} c_{ik}q_iq_k = \frac{1}{2}\sum_{j=1}^{n} \lambda_j\theta_j^2 \end{aligned}\right\} \tag{4}$$

The variables $\theta_1, \ldots, \theta_n$ are called *normal* or *principal coordinates*. The formulas (2) of transformation from arbitrary coordinates to normal coordinates may be given in vector notation as follows:

$$\boldsymbol{q} = \sum_{j=1}^{n} \theta_j \boldsymbol{u}_j \tag{5}$$

where

$$\boldsymbol{q} = \left\| \begin{matrix} q_1 \\ \cdot \\ \cdot \\ \cdot \\ q_n \end{matrix} \right\|, \quad \boldsymbol{u}_j = \left\| \begin{matrix} u_{1j} \\ \cdot \\ \cdot \\ \cdot \\ u_{nj} \end{matrix} \right\| \quad (j = 1, \ldots, n)$$

Since the transformation of coordinates (2) is nonsingular, the corresponding determinant is nonzero:

$$\det (u_{ij})^n_{i,\ j=1} \neq 0$$

i.e. the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ are linearly independent.

Utilizing the simple expressions (4) for T and Π in normal coordinates, form the Lagrange equations in these coordinates:

$$\ddot{\theta}_j + \lambda_j \theta_j = 0 \quad (j = 1, \ldots, n) \tag{6}$$

Each of these equations contains only one unknown function. The general solutions of the equations (6) are known to determine the harmonic oscillations

$$\theta_j = C_j \sin (\omega_j t + \alpha_j) \quad (j = 1, \ldots, n) \tag{7}$$

where C_j and α_j are arbitrary constants ($j = 1, \ldots, n$). Substituting these expressions into formula (5), we get a general formula for oscillations:

$$\boldsymbol{q} = \sum_{j=1}^{n} C_j \boldsymbol{u}_j \sin (\omega_j t + \alpha_j) \tag{8}$$

It is thus rigorously established that this formula embraces all small oscillations of a conservative system in the most general case.*

Assuming, in formula (8), all arbitrary constants, except C_j and α_j, equal to zero, we get the jth "principal" harmonic oscillation

$$\boldsymbol{q} = C_j \boldsymbol{u}_j \sin (\omega_j t + \alpha_j) \tag{9}$$

(In normal coordinates, this oscillation occurs when all $\theta_i = 0$ for $i \neq j$, and only the coordinate θ_j changes.) In the preceding section it was established that the square of the frequency $\lambda_j = \omega_j^2$ satisfies the frequency equation. Since for $\boldsymbol{q}$ there exist no harmonic oscillations of the type (9), other than those which occur as summands in the general formula (8), it follows that $\lambda_j = \omega_j^2$ ($j = 1, \ldots, n$) are all the roots of the secular equation. Besides, if some root is repeated p times, then to it there correspond p linearly independent amplitude vectors $\boldsymbol{u}_j$ determined from the system of linear equations (28) or (29) of the preceding section.

We have thus again proved that all the roots λ_j of the secular equation are real and positive and we have established that to the n frequencies $\omega_j = \sqrt{\lambda_j}$ there correspond n linearly independent amplitude vectors $\boldsymbol{u}_j$ ($j = 1, \ldots, n$).

---

[*] In the preceding section, this formula was established solely for the case when the secular equation has no multiple roots.

Substituting into the formula $A\ (q,\ q)$ the expression (5) for q in place of q, we obtain

$$A(q,\ q)=A\Big(\sum_{j=1}^{n}\theta_j u_j,\ \sum_{h=1}^{n}\theta_h u_h\Big)=\sum_{j,\ h=1}^{n}A(u_j,\ u_h)\,\theta_j\theta_h \tag{10}$$

On the other hand,

$$A(q,\ q)=\sum_{j=1}^{n}\theta_j^2 \tag{11}$$

Comparing (10) and (11), we get the following relations* for the amplitude vectors $u_j\ (j=1,\ \ldots,\ n)$, with the aid of which, using formula (5), the transformation to normal coordinates is accomplished:

$$A(u_i,\ u_j)=\delta_{ij}=\begin{cases}1, & i=j,\\ 0, & i\neq j,\end{cases}\quad (i,\ j=1,\ \ldots,\ n) \tag{12}$$

42. The Effect of Periodic External Forces on the Oscillations of a Conservative System

Aside from the potential forces $-\dfrac{\partial\Pi}{\partial q_i}$, let a system be acted upon by certain forces $Q_i=Q_i(t)\ (i=1,\ \ldots,\ n)$. Pass to the normal coordinates using the formulas

$$q_i=\sum_{j=1}^{n}u_{ij}\theta_j\quad (i=1,\ \ldots,\ n;\ \det(u_{ij})_{i,\ j=1}^{n}\neq 0) \tag{1}$$

To the forces Q_i in coordinates $q_i\ (i=1,\ \ldots,\ n)$ there correspond the forces Θ_j in coordinates $\theta_j\ (j=1,\ \ldots,\ n)$. Proceeding from the fact that the expressions for the elementary work of the forces are equal, let us establish a relation between Q_i and Θ_j

$$\sum_{i=1}^{n}Q_i\delta q_i=\sum_{j=1}^{n}\Theta_j\delta\theta_j \tag{2}$$

Noting that by virtue of formulas (1)

$$\delta q_i=\sum_{j=1}^{n}u_{ij}\delta\theta_j\quad (i=1,\ \ldots,\ n) \tag{3}$$

and substituting these expressions for δq_i into (2) we obtain

$$\sum_{j=1}^{n}\Big(\sum_{i=1}^{n}u_{ij}Q_i\Big)\delta\theta_j=\sum_{j=1}^{n}\Theta_j\delta\theta_j \tag{4}$$

---

* In other words, the vectors $u_j\ (j=1,\ \ldots,\ n)$ are orthonormal in the A metric (see footnote on page 207).

Hence, equating the coefficients of the independent increments of the normal coordinates $\delta\theta_j$, we get

$$\Theta_j = \sum_{i=1}^{n} u_{ij} Q_i \quad (j = 1, \ldots, n) \tag{5}$$

Thus, if the "old" coordinates q_i are expressed in terms of the "new" ones θ_j by means of the matrix $U = \| u_{ij} \|$:

$$\begin{aligned} q_1 &= u_{11}\theta_1 + u_{12}\theta_2 + \ldots + u_{1n}\theta_n, \\ q_2 &= u_{21}\theta_1 + u_{22}\theta_2 + \ldots + u_{2n}\theta_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ q_n &= u_{n1}\theta_1 + u_{n2}\theta_2 + \ldots + u_{nn}\theta_n \end{aligned} \quad (q = U\theta, \ \det U \neq 0) \tag{6}$$

then the "new" generalized forces Θ_j are expressed in terms of the "old" generalized forces Q_i by transposing the matrix U':

$$\begin{aligned} \Theta_1 &= u_{11}Q_1 + u_{21}Q_2 + \ldots + u_{n1}Q_n, \\ \Theta_2 &= u_{12}Q_1 + u_{22}Q_2 + \ldots + u_{n2}Q_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ \Theta_n &= u_{1n}Q_1 + u_{2n}Q_2 + \ldots + u_{nn}Q_n \end{aligned} \quad (\Theta = U'Q) \tag{7}$$

Comparing the matrix formulas $q = U\theta$ and $Q = (U')^{-1}\Theta$, we see that in passing from coordinates to forces the transformation matrix U is replaced* by the matrix $(U')^{-1}$.

This is expressed in words as follows: *the generalized forces are transformed contravariantly with respect to the coordinates.***

After learning how to determine the Θ_j from the given Q_i, let us write the Lagrange equations in normal coordinates using for T and Π the expressions (4) of the preceding section:

$$\ddot{\theta}_j + \omega_j^2\theta_j = \Theta_j(t) \quad (j = 1, \ldots, n) \tag{8}$$

Denote by θ_j^* $(j = 1, \ldots, n)$ an arbitrary particular solution of equation (8). Then the general solution of (8) will be

$$\theta_j = C_j \sin(\omega_j t + \alpha_j) + \theta_j^* \quad (j = 1, \ldots, n) \tag{9}$$

Let $\Theta_j(t)$ be the periodic force and let this force be sinusoidal with frequency Ω:

$$\Theta_j(t) = A_j \sin \Omega t \tag{10}$$

Then it will readily be seen that for the θ_j^* we can take

$$\theta_j^* = \frac{A_j}{\omega_j^2 - \Omega^2} \sin \Omega t \tag{11}$$

---

* If $U = \| u_{ij} \|_1^n$ is an orthogonal matrix, then $(U')^{-1} = U$, and the forces are transformed exactly like the coordinates.

** In the general case, when the transformation of coordinates is nonlinear, the generalized forces are transformed contravariantly with respect to the differentials of coordinates.

If $Q_i(t)$ $(i=1, \ldots, n)$, and hence also $\Theta_j(t)$ $(j=1, \ldots, n)$ are arbitrary periodic forces with period τ and frequency $\Omega=\frac{2\pi}{\tau}$, it follows that $\Theta_j(t)$ may be expanded in a Fourier series:

$$\Theta_j(t)=\sum_{m=0}^{\infty} A_{jm}\sin(m\Omega t+\varphi_{jm})$$
$$(j=1, \ldots, n) \tag{12}$$

Then, by virtue of the linearity of the equations (8),

$$\theta_j^*=\sum_{m=0}^{\infty}\frac{A_{jm}}{\omega_j^2-m^2\Omega^2}\sin(m\Omega t+\varphi_{jm}) \quad (j=1, \ldots, n) \tag{13}$$

If some $m\Omega$ coincides with ω_j, and the corresponding $A_{jm}\neq 0$, then for the coordinate θ_j we have the phenomenon of *resonance*.

Substituting expressions (9) for θ_j into the formula

$$\boldsymbol{q}=\sum_{j=1}^{n}\theta_j\boldsymbol{u}_j$$

we get

$$\boldsymbol{q}=\widehat{\boldsymbol{q}}+\boldsymbol{q}^* \tag{14}$$

where

$$\widehat{\boldsymbol{q}}=\sum_{j=1}^{n}C_j\boldsymbol{u}_j\sin(\omega_j t+\alpha_j) \tag{15}$$

are the *free oscillations* and

$$\boldsymbol{q}^*=\sum_{j=1}^{n}\theta_j^*\boldsymbol{u}_j$$

are the *forced oscillations* of the system, and $\boldsymbol{u}_j$ is the amplitude vector with coordinates $u_{1j}, u_{2j}, \ldots, u_{nj}$ $(j=1, \ldots, n)$.

43. Extremal Properties of Frequencies of a Conservative System. Rayleigh's Theorem on Frequency Variation with Change in Inertia and Rigidity of the System. Superimposition of Constraints

In Sec. 41 we considered the linear nonsingular transformation of coordinates

$$\boldsymbol{q}=\sum_{j=1}^{n}\theta_j\boldsymbol{u}_j \tag{1}$$

or, in scalar notation,

$$q_i=\sum_{j=1}^{n}u_{ij}\theta_j \quad (i=1, \ldots, n;\ \det(u_{ij})_{i,j=1}^{n}\neq 0) \tag{1'}$$

that accomplishes the transition to the normal coordinates $\theta_1, \ldots, \theta_n$ in which the quadratic forms*

$$A(q, q) = \sum_{i, k=1}^{n} a_{ik} q_i q_k, \quad C(q, q) = \sum_{i, k=1}^{n} c_{ik} q_i q_k \tag{2}$$

have the simple ("canonical") form:

$$A(q, q) = \sum_{j=1}^{n} \theta_j^2, \quad C(q, q) = \sum_{j=1}^{n} \omega_j^2 \theta_j^2 \tag{3}$$

From now on we shall assume that the principal oscillations are numbered in such fashion that their frequencies proceed in the following increasing order:

$$\omega_1 \leqslant \omega_2 \leqslant \ldots \leqslant \omega_n \tag{4}$$

Consider the ratio of quadratic forms (3)

$$\frac{C(q, q)}{A(q, q)} = \frac{\omega_1^2\theta_1^2 + \omega_2^2\theta_2^2 + \ldots + \omega_n^2\theta_n^2}{\theta_1^2 + \theta_2^2 + \ldots + \theta_n^2}. \tag{5}$$

for any $q \neq 0$ or, which is the same thing, for any values $\theta_1, \ldots, \theta_n$ that are not simultaneously equal to zero. Replacing all the ω_j^2 in the numerator of (5) by a smaller or equal number ω_1^2, we have

$$\frac{C(q, q)}{A(q, q)} \geqslant \omega_1^2 \tag{6}$$

On the other hand, from (5) it is at once obvious that for $\theta_2 = \ldots = \theta_n = 0$ the ratio $\frac{C(q, q)}{A(q, q)}$ attains the value ω_1^2. Hence,**

$$\omega_1^2 = \min \frac{C(q, q)}{A(q, q)} \tag{7}$$

---

* $C(q, q)$ is the doubled potential energy, and $A(q, q)$ is obtained from the expression for the doubled kinetic energy $A(\dot{q}, \dot{q})$ by replacing $\dot{q}$ by q.

** If on a real-number axis we plot the points $\omega_1^2, \omega_2^2, \ldots, \omega_n^2$ and at these points concentrate masses $m_1 = \theta_1^2$, $m_2 = \theta_2^2$, $\ldots$, $m_n = \theta_n^2$, then, by (5), the quantity $\frac{C(q, q)}{A(q, q)}$ will be the coordinate of the centre of these masses. Whence follow immediately the relations (6) and (7), since the centre of mass is always located between the extreme masses and coincides with one of the extreme masses when all the other masses are zero.

Now let a linear homogeneous constraint* be imposed on the system

$$l_1 q_1 + l_2 q_2 + \ldots + l_n q_n = 0 \quad \left(\sum_{i=1}^{n} l_i^2 > 0\right) \tag{8}$$

Expressing $q_1, q_2, \ldots, q_n$ in terms of the normal coordinates by means of the transformation (1), we will again have, in normal coordinates, a linear homogeneous constraint:

$$l_1' \theta_1 + l_2' \theta_2 + \ldots + l_n' \theta_n = 0 \quad \left(\sum_{j=1}^{n} l_j'^2 > 0\right) \tag{8'}$$

We abbreviate constraint (8) or (8′) to

$$L = 0$$

It is always possible to find the values θ_1 and θ_2 such that together with $\theta_3 = \ldots = \theta_n = 0$ they will satisfy the constraint equation (8′). According to formula (5), for the corresponding $\boldsymbol{q}$,

$$\frac{C(\boldsymbol{q}, \boldsymbol{q})}{A(\boldsymbol{q}, \boldsymbol{q})} = \frac{\omega_1^2 \theta_1^2 + \omega_2^2 \theta_2^2}{\theta_1^2 + \theta_2^2} \leqslant \omega_2^2$$

Therefore, also,**

$$\min_{L=0} \frac{C(\boldsymbol{q}, \boldsymbol{q})}{A(\boldsymbol{q}, \boldsymbol{q})} \leqslant \omega_2^2 \tag{9}$$

We shall now vary the constraint $L = 0$. Then the left side in the inequality (9) will vary remaining all the time smaller than or equal to ω_2^2. But for the constraint $\theta_1 = 0$ (here $l_1' = 1$, $l_2' = \ldots \ldots = l_n' = 0$) the ratio $\frac{C(\boldsymbol{q}, \boldsymbol{q})}{A(\boldsymbol{q}, \boldsymbol{q})}$ is given by the formula

$$\frac{C(\boldsymbol{q}, \boldsymbol{q})}{A(\boldsymbol{q}, \boldsymbol{q})} = \frac{\omega_2^2 \theta_2^2 + \ldots + \omega_n^2 \theta_n^2}{\theta_2^2 + \ldots + \theta_n^2}$$

and for this reason [by analogy with formulas (5) and (7)]

---

[*] If a nonlinear constraint is given and the equilibrium position satisfies the equation of the constraint, then there is no absolute term in power expansion of the left side of the constraint equation:

$$l_1 q_1 + l_2 q_2 + \ldots + l_n q_n + \sum_{i, k=1}^{n} l_{ik} q_i q_k + \ldots = 0$$

Besides, we assume that there are indeed linear terms, i.e. that $\sum_{i=1}^{n} l_i^2 > 0$. Then, rejecting second and higher-order terms, we represent the constraint equation as (8).

[**] The symbol on the left of the inequality (9) denotes the minimum of the ratio $\frac{C(\boldsymbol{q},\boldsymbol{q})}{A(\boldsymbol{q},\boldsymbol{q})}$, provided that only vectors $\boldsymbol{q} \neq 0$ are considered that satisfy the constraint equation (8).

$$\min_{\theta_1=0} \frac{C(q, q)}{A(q, q)} = \omega_2^2$$

Thus, of all the constraints of form $L=0$ the quantity

$$\min_{L=0} \frac{C(q, q)}{A(q, q)}$$

reaches its greatest value ω_2^2 with the constraint $\theta_1=0$. Hence,

$$\max \min_{L=0} \frac{C(q, q)}{A(q, q)} = \omega_2^2 \tag{10}$$

In place of one constraint $L=0$, it is possible to impose on the system a number of constraints $L_1=0, \ldots, L_{h-1}=0$. As was done in the special case of one constraint, it is possible to show that

$$\omega_h^2 = \max \min_{\substack{L_1=0 \\ \cdots \\ L_{h-1}=0}} \frac{C(q, q)}{A(q, q)} \quad (h=2, \ldots, n) \tag{11}$$

The formulas (7) and (11) express the extremal properties of frequencies of a conservative system. These properties are sometimes called *maximinimal* properties.

In place of the formulas (7) and (11), we can easily obtain similar formulas:

$$\omega_n^2 = \max \frac{C(q, q)}{A(q, q)} \tag{7'}$$

$$\omega_{n-h}^2 = \min \max_{\substack{L_1=0 \\ \cdots \\ L_h=0}} \frac{C(q, q)}{A(q, q)} \quad (h=1, \ldots, n-1) \tag{11'}$$

The extremal properties of the principal frequencies expressed by (7′) and (11′) are sometimes called *minimaximal* properties.

In addition to the given system, let us consider yet another conservative system with kinetic and potential energies

$$\frac{1}{2}\widetilde{A}(\dot{q}, \dot{q}) = \frac{1}{2}\sum_{i,\,k=1}^{n} \tilde{a}_{ik}\dot{q}_i\dot{q}_k, \quad \frac{1}{2}\widetilde{C}(q, q) = \frac{1}{2}\sum_{i,\,k=1}^{n} \tilde{c}_{ik}q_iq_k \tag{12}$$

and with principal frequencies

$$\tilde{\omega}_1 \leqslant \tilde{\omega}_2 \leqslant \ldots \leqslant \tilde{\omega}_n \tag{13}$$

For this system

$$\tilde{\omega}_1^2 = \min \frac{\tilde{C}(q, q)}{\tilde{A}(q, q)} \tag{14}$$

$$\tilde{\omega}_h^2 = \max_{\substack{L_1=0 \\ \cdots \\ L_{h-1}=0}} \min \frac{\tilde{C}(q, q)}{\tilde{A}(q, q)} \quad (h = 2, \ldots, n) \tag{15}$$

Let the new system have greater rigidity for the same inertia, i.e. for any q

$$\tilde{A}(q, q) = A(q, q), \quad \tilde{C}(q, q) \geqslant C(q, q)$$

or less inertia for the same rigidity

$$\tilde{A}(q, q) \leqslant A(q, q), \quad \tilde{C}(q, q) = C(q, q)$$

In both cases, for any $q \neq 0$,

$$\frac{C(q, q)}{A(q, q)} \leqslant \frac{\tilde{C}(q, q)}{\tilde{A}(q, q)} \tag{16}$$

But then both the minimums and the maximums of these ratios will be connected by the same inequality, i.e. from the inequality (16), by virtue of (7), (11), (14), and (15), there follows

$$\omega_j \leqslant \tilde{\omega}_j \quad (j = 1, \ldots, n) \tag{17}$$

Here, at least one of these relations has the sign $<$, so long as the identity*

$$\frac{C(q, q)}{A(q, q)} = \frac{\tilde{C}(q, q)}{\tilde{A}(q, q)} \tag{18}$$

does not hold.

We have arrived at the **Rayleigh theorem:****

*The principal frequencies increase when the rigidity of the system is increased or the inertia of the system is decreased.****

Let us find out how the imposition of constraints affects the magnitudes of the principal frequencies $\omega_1 \leqslant \omega_2 \leqslant \ldots \leqslant \omega_n$ of a conservative system.

---

[*] Indeed, according to (5), identity (18) holds in the case of $\omega_j = \tilde{\omega}_j$ $(j = 1, \ldots, n)$.

[**] This theorem was established by the English physicist Rayleigh in 1873 (Rayleigh J. W. *The Theory of Sound*. Vol. 1, 2nd ed. Macmillan, London, 1894).

[***] How much the principal oscillations increase may be judged in the first case by the difference $\tilde{C}(q, q) - C(q, q)$, and in the second case by the difference $\tilde{A}(q, q) - A(q, q)$ [see 7].

Impose on a system s independent linear constraints

$$\widehat{L}_1 = 0,\ \widehat{L}_2 = 0,\ \ldots,\ \widehat{L}_s = 0$$

Let the conservative system thus obtained with $n-s$ degrees of freedom have principal frequencies $\omega_1^* \leqslant \omega_2^* \leqslant \ldots \leqslant \omega_{n-s}^*$. Here,

$$\omega_1^{*2} = \min_{\substack{\widehat{L}_1=0\\ \cdots\\ \widehat{L}_s=0}} \frac{C(q,\ q)}{A(q,\ q)} \tag{19}$$

Comparing formula (19) with (7) and (11) (when $h-1=s$), we get

$$\omega_1 \leqslant \omega_1^* \leqslant \omega_{s+1} \tag{20}$$

In exactly the same way, for any $h \leqslant n-s$

$$\omega_h^{*2} = \max \min_{\substack{\widehat{L}_1=0,\ \ldots,\ \widehat{L}_s=0\\ L_1=0,\ \ldots,\ L_{h-1}=0}} \frac{C(q,\ q)}{A(q,\ q)} \tag{21}$$

Here the constraints $\widehat{L}_1=0,\ \ldots,\ \widehat{L}_s=0$ are fixed, while the constraints $L_1=0,\ \ldots,\ L_{h-1}=0$ vary. Comparing (21) with (11) and with the formula

$$\omega_{h+s}^2 = \max \min_{\substack{L_1=0\\ \cdots\cdots\\ L_{h+s-1}=0}} \frac{C(q,\ q)}{A(q,\ q)}$$

in which all $s+h-1$ constraints vary, we will have

$$\omega_h \leqslant \omega_h^* \leqslant \omega_{h+s} \quad (h=1,\ \ldots,\ n-s) \tag{22}$$

Formulas (22) show that in the imposition of s independent constraints each of the first $n-s$ principal frequencies increases and in the process does not exceed the old principal frequency, the number of which is s units greater than the number of the given frequency.

1. As an application of this proposition, it may be shown that the roots $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$ of the secular equation $\Delta(\lambda) \equiv \det(c_{ik} - \lambda a_{ik})_{i,\ k=1}^{n} = 0$ are separated by the roots $\lambda_1^* \leqslant \ldots \leqslant \lambda_{n-1}^*$ of the equation[*]

$$\Delta_1(\lambda) = \det(c_{ik} - \lambda a_{ik})_{i,\ k=1}^{n-1} = 0$$

i.e.

$$\lambda_1 \leqslant \lambda_1^* \leqslant \lambda_2 \leqslant \lambda_2^* \leqslant \ldots \leqslant \lambda_{n-1}^* \leqslant \lambda_n \tag{23}$$

Indeed, the equation $\Delta_1(\lambda) = 0$ is a secular equation for a conservative system obtained from the initial one by superimposition of one constraint

---

[*] $\Delta_1(\lambda)$ is the principal minor of order $(n-1)$ in the determinant $\Delta(\lambda)$. It is sometimes said that inequalities (23) express the "separation theorem" for roots of the secular equation. Inequalities (23) may be used to find the lower and upper bounds of the roots of a secular equation (see for example [2]).

$q_n = 0$. Therefore, setting $\lambda_k = \omega_k^2$ ($k = 1, \ldots, n$), $\lambda_j^* = \omega_j^{*2}$ ($j = 1, \ldots, n - 1$), we at once get the inequalities (23) from the inequalities (22) for $s = 1$.

2. We indicate yet another noteworthy application of the proposition on variation of frequencies upon the superimposition of constraints. Cracks in a glass are known to be revealed by tapping one's fingers on it. The reason lies in the fact that a glass without cracks has supplementary constraints (compared with a glass with cracks) between its parts. That is why in a glass without cracks the oscillations must be higher.

44. Small Oscillations of Elastic Systems

As an important example of small oscillations of a conservative system, consider n masses $m_1, m_2, \ldots, m_n$ concentrated at n points (1), (2), ..., (n) of an elastic system S (a string, rod, membrane, plate, and the like) having finite dimensions and in some way attached at the ends. We shall assume that the displacements (deflections) $y_1, y_2, \ldots, y_n$ of points (1), (2), ..., (n) of the system S and the forces $F_1, F_2, \ldots, F_n$ acting on the masses $m_1, m_2, \ldots, m_n$ are parallel to one and the same direction and for this reason are determined by their algebraic magnitudes (Fig. 50). The deflections $y_1, y_2, \ldots, y_n$ may be regarded as independent coordinates of the system, and the forces $F_1, F_2, \ldots, F_n$ as the corresponding generalized forces, since the elementary work of these forces is equal to

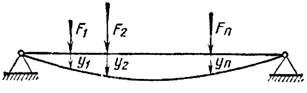

Fig. 50

$$\sum_{i=1}^{n} F_i \, dy_i$$

When investigating free oscillations, we take for the forces $F_1, F_2, \ldots, F_n$ the elastic forces $F_1^*, \ldots, F_n^*$ acting on the masses $m_1, m_2, \ldots, m_n$ of the elastic system S, given the deflections $y_1, y_2, \ldots, y_n$. The system under consideration of n particles acted upon by elastic forces is a conservative system and has a definite potential energy $\Pi(y_1, y_2, \ldots, y_n)$. Expanding the function $\Pi(y_1, y_2, \ldots, y_n)$ in a power series and retaining only the quadratic terms (see Sec. 40), we get for Π the expression

$$\Pi = \frac{1}{2} \sum_{i,\,k=1}^{n} c_{ik} y_i y_k \quad (c_{ik} = c_{ki}, \ i, k = 1, \ldots, n) \tag{1}$$

Then for the elastic forces $F_i^* = -\frac{\partial \Pi}{\partial y_i}$ $(i = 1, \ldots, n)$ we find

$$F_i^* = -\sum_{k=1}^{n} c_{ik} y_k \quad (i = 1, \ldots, n) \tag{2}$$

Considering the position of equilibrium $y_1 = \ldots = y_n = 0$ to be stable, we assume that the quadratic form (1) that expresses the potential energy as a function of deflection is positive definite:

$$\sum_{i,\,k=1}^{n} c_{ik} y_i y_k > 0 \quad \left(\sum_{i=1}^{n} y_i^2 > 0\right) \tag{3}$$

The kinetic energy of the system has a simple form:

$$T = \sum_{i=1}^{n} \frac{1}{2} m_i \dot{y}_i^2 \tag{4}$$

In the search for the harmonic oscillation $y_i = u_i \sin(\omega t + \alpha)$, we arrive (as we did in Sec. 40) at the frequency equation

$$\begin{vmatrix} c_{11} - m_1\lambda & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} - m_2\lambda & \ldots & c_{2n} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ c_{n1} & c_{n2} & \ldots & c_{nn} - m_n\lambda \end{vmatrix} = 0 \quad (\lambda = \omega^2) \tag{5}$$

and at the algebraic equations for determining the amplitudes:

$$\sum_{k=1}^{n} (c_{ik} - \lambda m_i \delta_{ik}) u_k = 0 \quad (i = 1, \ldots, n) \tag{6}$$

The system has n frequencies

$$\omega_1 \leqslant \omega_2 \leqslant \ldots \leqslant \omega_n \tag{7}$$

and the corresponding amplitude columns $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$; the free oscillations are determined by the formula

$$\boldsymbol{y} = \sum_{j=1}^{n} C_j \boldsymbol{u}_j \sin(\omega_j t + \alpha_j) \tag{8}$$

where C_j and α_j $(j = 1, \ldots, n)$ are arbitrary constants determined from the initial conditions.

Let external forces $F_1, \ldots, F_n$ give rise to static deflection $y_1, \ldots, y_n$. Then the forces F_i are balanced by the elastic forces F_i^* $(F_i = -F_i^*;\ i = 1, \ldots, n)$ and for this reason, according to equation (2),

$$F_i = \sum_{k=1}^{n} c_{ik} y_k \quad (i = 1, \ldots, n) \tag{9}$$

An important role in the study of elastic systems is played by the matrix $G = \| g_{ik} \|_1^n$, which is the inverse* of the matrix $C = \| c_{ik} \|_1^n$:

$$G = C^{-1}$$

Using matrix G it is possible to solve the system of relations (9) for $y_1, \ldots, y_n$ and to represent it in the form

$$y_i = \sum_{k=1}^{n} g_{ik} F_k \quad (i = 1, \ldots, n) \tag{10}$$

The quantity g_{ik} is equal to the deflection at the point (i) caused by a unit external force applied at the point (k), and is called the influence coefficient of point (k) on the point (i) $(i, k = 1, \ldots, n)$. From the symmetricalness of the matrix C there follows the symmetricalness of the inverse matrix G composed of the influence coefficients**

$$g_{ik} = g_{ki} \quad (i, k = 1, \ldots, n) \tag{11}$$

and from the positive definiteness of the forms (3) there follows the positive definiteness of the quadratic form

$$G(F, F) = \sum_{i, k=1}^{n} g_{ik} F_i F_k > 0 \quad \left(\sum_{i=1}^{n} F_i^2 > 0\right) \tag{12}$$

since the quadratic form $\sum\limits_{i, k=1}^{n} c_{ik} y_i y_k$ goes into the form (12) in the case of transformation of variables (10):

$$\Pi = \frac{1}{2} \sum_{i, k=1}^{n} c_{ik} y_i y_k = \frac{1}{2} \sum_{i=1}^{n} F_i y_i = \frac{1}{2} \sum_{i, k=1}^{n} g_{ik} F_i F_k \tag{13}$$

Consider a *linear* elastic system S—a string or a rod with ends fixed as usual. It may be shown that in this case the matrix of the coefficients of influence, G, possesses the following properties:

1°. All minors (and not only the principal ones!) of any order of the matrix G are nonnegative:

$$\begin{vmatrix} g_{i_1 k_1} & \cdots & g_{i_1 k_p} \\ \cdots & \cdots & \cdots \\ g_{i_p k_1} & \cdots & g_{i_p k_p} \end{vmatrix} \geqslant 0$$

$$(0 \leqslant i_1 \leqslant \ldots \leqslant i_p \leqslant n;\ 0 \leqslant k_1 \leqslant \ldots \leqslant k_p \leqslant n;\ p = 1, \ldots, n)$$

---

* The matrix C is not singular (det $C \neq 0$), since the quadratic form (3) is positive definite.

** Equation (11) expresses the so-called Maxwell principle of reciprocity: "a deflection at point (i) due to a unit force applied to point (k) is equal to the deflection at (k) under a unit force applied to point (i)".

2°. $g_{ik} > 0$ for $|i - k| \leqslant 1$ $(i, k = 1, \ldots, n)$.

3°. The determinant $\det G = |g_{ik}|_1^n > 0$.

Matrices having the properties 1°, 2° and 3° are called *oscillatory matrices.*

It will be noted that for every positive definite matrix G, the property 3° and also the inequalities 1° for the principal minors and the inequalities 2° for the diagonal elements g_{ii} $(i = 1, \ldots, n)$ are valid. However, the nonnegativity of the nonprincipal minors of any order p* and the positivity of the elements $g_{12}, \ldots, g_{n-1,\,n}$ are specific properties of the matrix of influence coefficients of a linear elastic system.

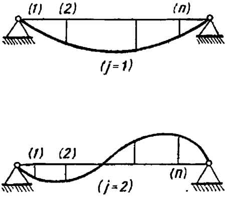

Fig. 51

From the oscillatory nature of the matrix of influence coefficients follow the basic "oscillatory" properties of elastic oscillations of a linear system:

1°. All frequencies are distinct:

$$\omega_1 < \omega_2 < \ldots < \omega_n$$

2°. All the amplitudes $u_{11}, u_{21}, \ldots, u_{n1}$ in the first principal oscillation (with frequency ω_1) are different from zero and have the same signs.

3°. Among the amplitudes $u_{1j}, u_{2j}, \ldots, u_{nj}$ in the jth principal oscillation (with frequency ω_j) there are exactly $j - 1$ changes of sign $(j = 1, 2, \ldots, n)$ (Fig. 51).

An investigation of oscillatory matrices and a substantiation of oscillatory properties of elastic oscillations go beyond the scope of this book.**

Example. Consider the classical problem of the oscillation of a string of length l with attached ends for the case when the entire mass of the string is concentrated in n equidistant (from one another and from the endpoints) points, the concentrated masses being all equal (and equal to m) (Fig. 52).

---

[*] In particular the nonnegativity of the off-diagonal elements of the matrix G ($g_{ik} \geqslant 0$ when $i \neq k$).

[**] This material can be found in [7].

The elongation of the ith segment (between points with deflection y_i and deflection y_{i+1}) will be expressed (to within fourth-order infinitesimals) as follows:

$$\sqrt{\left(\frac{l}{n+1}\right)^2+(y_{i+1}-y_i)^2}-\frac{l}{n+1}=\frac{l}{n+1}\left(\sqrt{1+\left(\frac{n+1}{l}\right)^2(y_{i+1}-y_i)^2}-1\right)\approx$$
$$\approx\frac{n+1}{2l}(y_{i+1}-y_i)^2$$

Taking the tension σ of the string to be constant*, we get an expression for the potential energy:

$$\Pi=\frac{c}{2}\sum_{i=0}^{n}(y_{i+1}-y_i)^2 \quad \left(y_0=y_{n+1}=0;\quad c=\frac{\sigma(n+1)}{l}\right) \tag{14}$$

The kinetic energy is of the simple form

$$T=\frac{m}{2}\sum_{i=1}^{n}\dot{y}_i^2 \tag{15}$$

We choose an indirect path to find the principal frequencies and the corresponding amplitude vectors. Write equations (6) for the amplitudes, making use of the expressions (14) and (15) for Π and T. Divide each of the equations (6) termwise by c and introduce the abridged notation:

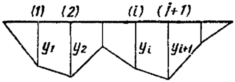

Fig. 52

$$1-\frac{m}{2c}\omega^2=\cos\theta \tag{16}$$

where θ is an auxiliary quantity. Then equations (6) will take the following form for the amplitudes:

$$u_{k-1}-2u_k\cos\theta+u_{k+1}=0\ (k=1,\ \ldots,\ n) \tag{17}$$

where

$$u_0=u_{n+1}=0 \tag{18}$$

The algebraic equations (6) may be satisfied by putting**

$$u_k=\sin k\theta\ (k=0,\ 1,\ \ldots,\ n+1) \tag{19}$$

---

* This supposition is justified in that only small deflections $y_1, y_2, \ldots, y_n$ are considered.

** Substituting (19) into (17) we get the trigonometric identities $\sin(k-1)\theta-2\sin k\theta\cos\theta+\sin(k+1)\theta=0$ $(k=0,\ 1,\ \ldots,\ n)$

Here, each of the "boundary" conditions (18) is satisfied automatically, and the second one yields the condition for determining the desired frequencies:

$$\sin(n+1)\theta=0 \tag{20}$$

Hence, $\theta_j=\frac{j\pi}{n+1}$ $(j=1, \ldots, n)$ and, by (16),

$$\omega_j^2=\frac{2c}{m}(1-\cos\theta_j)$$

i.e.

$$\left.\begin{array}{c}\omega_j=2\sqrt{\frac{c}{m}}\sin\frac{\theta_j}{2}=2\sqrt{\frac{c}{m}}\sin\frac{j\pi}{2(n+1)}\\ \left(j=1, \ldots, n;\ c=\frac{(n+1)\sigma}{l}\right)\end{array}\right\} \tag{21}$$

To find the amplitudes of the jth principal oscillation $u_{1j}, u_{2j}, \ldots, u_{nj}$ we put $\theta=\theta_j$ in equations (19):

$$u_{kj}=\sin k\theta_j=\sin\frac{kj\pi}{n+1}\quad (k,\ j=1, \ldots, n) \tag{22}$$

An arbitrary free oscillation of the system is determined by the formula

$$y_k=\sum_{j=1}^{n} C_j u_{kj}\sin(\omega_j t+\alpha_j)=\sum_{j=1}^{n} C_j\sin\frac{kj\pi}{n+1}\sin\left(2\sin\frac{j\pi}{2(n+1)}\sqrt{\frac{c}{m}}\,t+\alpha_j\right) \tag{23}$$

From (21) and (22) it is at once evident that the principal oscillations obtained possess the oscillatory properties 1° to 3°.

Lagrange demonstrated how it is possible, using these formulas and a limit passage, to obtain the free oscillations of a homogeneous string (with fixed ends), the mass of which is no longer concentrated at n points but is distributed uniformly along the string, which has density ρ.

In this problem we put $m=\frac{\rho l}{n}$, and then find the discrete analogue of a homogeneous string with principal frequencies

$$\omega_j^{(n)}=\frac{2}{l}\sqrt{\frac{\sigma}{\rho}n(n+1)}\sin\frac{j\pi}{2(n+1)}\quad (j=1, \ldots, n) \tag{24}$$

In the limit, as $n\to\infty$, we get the following familiar expressions for the frequencies ω_j of a fixed homogeneous string:

$$\omega_j=j\frac{\pi}{l}\sqrt{\frac{\sigma}{\rho}}\quad (j=1, \ldots, n) \tag{25}$$

These formulas express the Mersenne law, according to which *all frequencies are integral multiples of the frequency of the fundamental tone* $\omega_1=\frac{\pi}{l}\sqrt{\frac{\sigma}{\rho}}$ *and each of the frequencies is directly proportional to the square root of the tension and inversely proportional to the length and the square root of the density.*

Give the jth harmonic oscillation of a homogeneous string in the form

$$y_j(x, t) = u_j(x) \sin(\omega_j t + \alpha_j) \quad (0 \leqslant x \leqslant l) \tag{26}$$

where $u_j(x)$ is the amplitude deflection in this oscillation.

Considering that the amplitude deflection $u_j(x)$ may be obtained from the quantities (22) by passage to the limit

$$u_j(x) = \lim u_{kj}$$

as n tends to infinity and $\frac{kl}{n+l} \to x$, we get from (22)

$$u_j(x) = \sin \frac{j\pi x}{l} \quad (j = 1, 2, \ldots)$$

Then the free oscillation of a homogeneous string, which is obtained by linear superposition of the principal oscillations (26), is expressed by the formula

$$y(x, t) = \sum_{j=1}^{\infty} C_j \sin \frac{j\pi x}{l} \sin(\omega_j t + \alpha_j)$$

where C_j and α_j $(j = 1, 2, \ldots)$ are arbitrary constants.

45. Small Oscillations of a Scleronomic System under the Action of Forces Not Explicitly Dependent on the Time

Write the Lagrange equations for a scleronomic system for the case when the generalized forces Q_i depend solely on the coordinates and the velocities:

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = Q_i(q_k, \dot{q}_k) \quad (i = 1, \ldots, n) \tag{1}$$

Let the coordinate origin be a position of equilibrium. Then (see Sec. 40) the kinetic energy, to within third-order infinitesimals in q_i and $\dot{q}_i$ $(i = 1, \ldots, n)$, may be represented by a quadratic form with constant coefficients:

$$T = \frac{1}{2} \sum_{i, k=1}^{n} a_{ik} \dot{q}_i \dot{q}_k \tag{2}$$

where $a_{ik} = a_{ki}$ $(i, k = 1, \ldots, n)$.

Let us now expand the generalized forces $Q_i(q_k, \dot{q}_k)$ in a power series of q_k and $\dot{q}_k$

$$Q_i = Q_{i0} + \sum_{k=1}^{n} \left[\left(\frac{\partial Q_i}{\partial q_k} \right)_0 q_k + \left(\frac{\partial Q_i}{\partial \dot{q}_k} \right)_0 \dot{q}_h \right] + (**) \tag{3}$$

Since the coordinate origin is a position of equilibrium, all the generalized forces, for zero coordinates and velocities must be zero, i.e.

$$Q_{i0}=0 \quad (i=1, \ldots, n) \tag{4}$$

We introduce the following notation:

$$b_{ik}=-\left(\frac{\partial Q_i}{\partial \dot{q}_k}\right)_0, \quad c_{ik}=-\left(\frac{\partial Q_i}{\partial q_k}\right)_0 \quad (i, k=1, \ldots, n) \tag{5}$$

Then, rejecting in the expansion (3) all terms of second and higher orders of smallness, we will have

$$Q_i=-\sum_{k=1}^{n}(b_{ik}\dot{q}_k+c_{ik}q_k) \quad (i=1, \ldots, n) \tag{6}$$

Substituting into the Lagrange equations (1) the expressions (2) and (6) for the kinetic energy and for the generalized forces, we get the linear differential equations of motion for small oscillations of a scleronomic system:

$$\sum_{k=1}^{n}(a_{ik}\ddot{q}_k+b_{ik}\dot{q}_k+c_{ik}q_k)=0 \quad (i=1, \ldots, n) \tag{7}$$

Denote by $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ the square matrices*

$$\boldsymbol{A}=\|a_{ik}\|_1^n, \quad \boldsymbol{B}=\|b_{ik}\|_1^n, \quad \boldsymbol{C}=\|c_{ik}\|_1^n$$

and by $\boldsymbol{q}$ the column made up of $q_1, \ldots, q_n$. Then in matrix notation, the system of differential equations (7) will look like

$$\boldsymbol{A}\ddot{\boldsymbol{q}}+\boldsymbol{B}\dot{\boldsymbol{q}}+\boldsymbol{C}\boldsymbol{q}=0 \tag{8}$$

We shall seek the solution of (8) of the form

$$\boldsymbol{q}=\boldsymbol{u}e^{\mu t} \tag{9}$$

where $\boldsymbol{u}$ is a column with constant elements $u_1, \ldots, u_n$, and μ is a number.

Substituting (9) into the matrix equation (8) and cancelling $e^{\mu t}$, we get

$$(\boldsymbol{A}\mu^2+\boldsymbol{B}\mu+\boldsymbol{C})\,\boldsymbol{u}=0 \tag{10}$$

or, in expanded form,

$$\sum_{k=1}^{n}(a_{ik}\mu^2+b_{ik}\mu+c_{ik})\,u_k=0 \quad (i=1, \ldots, n) \tag{10'}$$

---

* Note that $\boldsymbol{A}$ is a positive definite symmetric matrix (this fact is not utilized here).

For the system (10) or (10′) to have a nonzero solution $\boldsymbol{u}$, it is necessary and sufficient that the determinant of the system be equal to zero:

$$\Delta(\mu) \equiv \det(\boldsymbol{A}\mu^2 + \boldsymbol{B}\mu + \boldsymbol{C}) = 0 \tag{11}$$

or in the expanded form,

$$\Delta(\mu) \equiv \begin{vmatrix} a_{11}\mu^2 + b_{11}\mu + c_{11} & \dots & a_{1n}\mu^2 + b_{1n}\mu + c_{1n} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ a_{n1}\mu^2 + b_{n1}\mu + c_{n1} & \dots & a_{nn}\mu^2 + b_{nn}\mu + c_{nn} \end{vmatrix} = 0 \tag{11′}$$

Equation (11) is called the *secular equation* of the given system. This is an algebraic equation in μ of degree $2n$.

We confine ourselves to the basic case when all the roots of the secular equation $\mu_1, \dots, \mu_{2n}$ are distinct. To each root μ_h there corresponds a certain nonzero solution $\boldsymbol{u}_h \equiv (u_{1h}, \dots, u_{nh})$ of the system of homogeneous algebraic equations (10) and, hence, a particular solution $\boldsymbol{u}_h e^{\mu_h t}$ of the system of differential equations (8) ($h = 1, \dots, 2n$). The general solution of this system is obtained as a linear combination (with arbitrary constant coefficients) of these particular solutions:

$$\boldsymbol{q} = \sum_{h=1}^{2n} C_h \boldsymbol{u}_h e^{\mu_h t} \tag{12}$$

Of particular importance is the case when all real parts of the roots μ_h are negative:

$$\operatorname{Re}\mu_h < 0 \quad (h = 1, \dots, 2n)$$

In this case, the position of equilibrium of the system is asymptotically stable not only for the linearized system (8) but also for the initial nonlinear scleronomic system with the differential equations (1) (see Sec. 38).

In conclusion, note that for a conservative system, $\boldsymbol{B} = \| b_{ik} \|_1^n = 0$, whereas, $\boldsymbol{A} = \| a_{ik} \|_1^n$ and $\boldsymbol{C} = \| c_{ik} \|_1^n$ are symmetric positive-definite matrices. The secular equation $\det(\boldsymbol{A}\mu^2 + \boldsymbol{C}) = 0$ passes into the equation $\det(\boldsymbol{C} - \lambda\boldsymbol{A}) = 0$ from Sec. 40, if we put $\mu = i\sqrt{\lambda}$ $(i = \sqrt{-1})$. But, as was shown in Sec. 40, the equation $\det(\boldsymbol{C} - \lambda\boldsymbol{A}) = 0$ has only positive and real roots. Therefore, in the case of a conservative system, equation (11) has only imaginary roots.

46. Rayleigh's Dissipative Function. The Effect of Small Dissipative Forces on the Oscillations of a Conservative System

Consider an important special case when the asymptotic stability of a position of equilibrium is predetermined and there is no necessity to resort to the stability criterion given in Sec. 39.

In the expressions for the generalized forces [see equation (6) on page 229]

$$Q_i = -\sum_{k=1}^{n} b_{ik}\dot{q}_k - \sum_{k=1}^{n} c_{ik}q_k \quad (i = 1, \ldots, n) \tag{1}$$

let the matrices $B = \| b_{ik} \|_1^n$ and $C = \| c_{ik} \|_1^n$, which are composed of the coefficients, be symmetric and positive definite.

Then, by introducing the potential energy Π and Rayleigh's dissipative function R (see Sec. 8), which are given by the positive-definite quadratic forms

$$\Pi = \frac{1}{2} \sum_{i,\,k=1}^{n} c_{ik}q_iq_k > 0, \quad R = \frac{1}{2} \sum_{i,\,k=1}^{n} b_{ik}\dot{q}_i\dot{q}_k > 0$$
$$\left(\sum_{i=1}^{n} q_i^2 > 0, \quad \sum_{i=1}^{n} \dot{q}_i^2 > 0\right) \tag{2}$$

we can rewrite formulas (1) as

$$Q_i = -\frac{\partial \Pi}{\partial q_i} - \frac{\partial R}{\partial \dot{q}_i} \quad (i = 1, \ldots, n) \tag{3}$$

Aside from the potential forces $-\frac{\partial \Pi}{\partial q_i}$ $(i = 1, \ldots, n)$, the system is acted upon by the dissipative forces $-\frac{\partial R}{\partial \dot{q}_i}$ defined by the Rayleigh function. In Sec. 8 it was explained that in this case the system is definite dissipative. Since, according to the first of the formulas (2), the potential energy in an equilibrium position has a strict minimum and the equilibrium position is isolated, it follows that (see the theorem on page 178) the equilibrium position is asymptotically stable.

Thus, the dissipative forces defined by the Rayleigh function not only do not disturb the stability of the equilibrium position of a conservative system, but even make this position (in certain cases) asymptotically stable.

In the case under consideration it is possible to establish simple formulas to evaluate the roots of the secular equation. Let us again seek a solution of the form $ue^{\mu t}$. To determine the column u, we get the equation (see page 229)

$$(A\mu^2 + B\mu + C)\,u = 0 \tag{4}$$

or, in expanded form,

$$\sum_{k=1}^{n} (a_{ik}\mu^2 + b_{ik}\mu + c_{ik})\, u_k = 0 \qquad (i = 1, \ldots, n) \tag{4'}$$

Multiplying both parts of the *i*th equation of (4′) by $\bar{u}_i$ ($\bar{u}_i$ is the complex conjugate of u_i) and summing over i, we find

$$\mu^2 \sum_{i,k=1}^{n} a_{ik}\bar{u}_i u_k + \mu \sum_{i,k=1}^{n} b_{ik}\bar{u}_i u_k + \sum_{i,k=1}^{n} c_{ik}\bar{u}_i u_k = 0$$

or in abridged notation

$$A(\boldsymbol{u}, \bar{\boldsymbol{u}})\,\mu^2 + B(\boldsymbol{u}, \bar{\boldsymbol{u}})\,\mu + C(\boldsymbol{u}, \bar{\boldsymbol{u}}) = 0 \tag{5}$$

where $A(\boldsymbol{u}, \bar{\boldsymbol{u}}) > 0$, $B(\boldsymbol{u}, \bar{\boldsymbol{u}}) > 0$ and $C(\boldsymbol{u}, \bar{\boldsymbol{u}}) > 0$.*

Thus, any root μ of the secular equation satisfies the quadratic equation (5) with positive coefficients. From this it follows immediately that $\operatorname{Re}\mu < 0$.

If the secular equation has the complex root $\mu = \gamma + i\delta$, then this same equation also has the complex conjugate root $\bar{\mu} = \gamma - i\delta$. The numbers μ and $\bar{\mu}$ are roots of the quadratic equation (5). Therefore, assuming $\boldsymbol{u} = \boldsymbol{v} + i\boldsymbol{w}$, $\bar{\boldsymbol{u}} = \boldsymbol{v} - i\boldsymbol{w}$ ($\boldsymbol{v}$, $\boldsymbol{w}$ are real column vectors), we can write **

$$\left.\begin{aligned} 2\operatorname{Re}\mu = \mu + \bar{\mu} = -\frac{B(\boldsymbol{u}, \bar{\boldsymbol{u}})}{A(\boldsymbol{u}, \bar{\boldsymbol{u}})} = -\frac{B(v, v) + B(w, w)}{A(v, v) + A(w, w)} < 0 \\ |\mu|^2 = \mu\bar{\mu} = \frac{C(\boldsymbol{u}, \bar{\boldsymbol{u}})}{A(\boldsymbol{u}, \bar{\boldsymbol{u}})} = \frac{C(v, v) + C(w, w)}{A(v, v) + A(w, w)} \end{aligned}\right\} \tag{6}$$

To the complex conjugate roots μ and $\bar{\mu}'$ there correspond the complex conjugate oscillations $\boldsymbol{u}e^{\mu t}$ and $\bar{\boldsymbol{u}}e^{\bar{\mu} t}$. It is possible to reduce the sum of the appropriate terms in the expression for $\boldsymbol{q}$ [see formula (12) on page 230] to the real form with complex conjugate values of the arbitrary constants $C = \frac{1}{2}(F + iG)$, $\bar{C} = \frac{1}{2}(F - iG)$:

$$\begin{aligned} C\boldsymbol{u}e^{\mu t} + \bar{C}\bar{\boldsymbol{u}}e^{\bar{\mu} t} &= \frac{1}{2}(F + iG)(\boldsymbol{v} + i\boldsymbol{w})\, e^{(\gamma + i\delta)t} + \\ &\quad + \frac{1}{2}(F - iG)(\boldsymbol{v} - i\boldsymbol{w})\, e^{(\gamma - i\delta)t} = \\ &= \operatorname{Re}\{[F\boldsymbol{v} - G\boldsymbol{w} + i(F\boldsymbol{w} + G\boldsymbol{v})]\, e^{\gamma t}(\cos\delta t + i\sin\delta t)\} = \\ &= e^{\gamma t}[(F\boldsymbol{v} - G\boldsymbol{w})\cos\delta t - (F\boldsymbol{w} + G\boldsymbol{v})\sin\delta t] \end{aligned} \tag{7}$$

---

* See 5° on page 206.

** See formula (20) on page 206.

If we have two real roots μ and μ' and if we denote the corresponding columns by $\boldsymbol{u}$ and $\boldsymbol{u}'$, then by multiplying both sides of the ith equation (4′) by u_i' (in place of $\overline{u}_i$), we get the following equation in place of (5):

$$A(\boldsymbol{u}, \boldsymbol{u}')\mu^2 + B(\boldsymbol{u}, \boldsymbol{u}')\mu + C(\boldsymbol{u}, \boldsymbol{u}') = 0 \tag{8}$$

Interchanging the roles of the vectors $\boldsymbol{u}$ and $\boldsymbol{u}'$, we conclude that the number μ' also satisfies the equation (8). Therefore

$$\mu + \mu' = -\frac{B(\boldsymbol{u}, \boldsymbol{u}')}{A(\boldsymbol{u}, \boldsymbol{u}')}, \quad \mu\mu' = \frac{C(\boldsymbol{u}, \boldsymbol{u}')}{A(\boldsymbol{u}, \boldsymbol{u}')} \tag{9}$$

Let us now see how the principal oscillations of conservative systems change under the action of small dissipative forces [see 27]. We introduce the normal coordinates $\theta_1, \ldots, \theta_n$. In these coordinates

$$T = \frac{1}{2}\sum_{i=1}^{n} \dot{\theta}_i^2, \quad \Pi = \frac{1}{2}\sum_{i=1}^{n} \omega_i^2\theta_i^2 \tag{10}$$

$$R = \frac{1}{2}\sum_{i=1}^{n} \beta_i\dot{\theta}_i^2 + \frac{1}{2}\sum_{\substack{i,k=1\\(i\neq k)}}^{n} \beta_{ik}\dot{\theta}_i\dot{\theta}_k \tag{11}$$

where ω_i $(i = 1, \ldots, n)$ are the principal frequencies of the conservative system, and the coefficients in the expression (11) for the Rayleigh functions β_i, β_{ik} $(i, k = 1, \ldots, n;\ \beta_{ik} = \beta_{ki}$ for $i \neq k)$ are small (squares and products of these quantities can be ignored). From the positive definiteness of the quadratic form (11) it follows that $\beta_i > 0$ $(i = 1, \ldots, n)$.

We form the Lagrange equations

$$\ddot{\theta}_i + \beta_i\dot{\theta}_i + \omega_i^2\theta_i + \sum_{\substack{k=1\\(k\neq i)}}^{n} \beta_{ik}\dot{\theta}_k = 0 \quad (i = 1, \ldots, n) \tag{12}$$

Then, substituting

$$\theta_i = \chi_i e^{\mu t} \quad (i = 1, \ldots, n) \tag{13}$$

and cancelling $e^{\mu t}$, we get the following system of linear equations:

$$(\mu^2 + \beta_i\mu + \omega_i^2)\chi_i + \mu\sum_{\substack{k=1\\(k\neq i)}}^{n} \beta_{ik}\chi_k = 0 \quad (i = 1, \ldots, n) \tag{14}$$

Equating the determinant of this system to zero, we find the secular equation:

$$\begin{vmatrix} \mu^2+\beta_1\mu+\omega_1^2 & \beta_{12}\mu & \dots & \beta_{1n}\mu \\ \beta_{21}\mu & \mu^2+\beta_2\mu+\omega_2^2 & \dots & \beta_{2n}\mu \\ \dots & \dots & \dots & \dots \\ \beta_{n1}\mu & \beta_{n2}\mu & \dots & \mu^2+\beta_n\mu+\omega_n^2 \end{vmatrix} = 0 \qquad (15)$$

Expanding the determinant and neglecting terms containing products of the small coefficients β_i, β_{ik}, we can represent the secular equation in the following form:

$$\prod_{k=1}^{n} (\mu^2+\beta_k\mu+\omega_k^2) = 0 \qquad (16)$$

To a first approximation, the roots of the secular equation are of the form

$$\mu_k \approx -\frac{\beta_k}{2} \pm i\omega_k \qquad (k=1, \dots, n) \qquad (17)$$

Let us find the amplitudes $\chi_1, \chi_2, \dots, \chi_n$ for $\mu=\mu_1=-\frac{\beta_1}{2}+i\omega_1$. Substituting this value of μ into the coefficients of the last $n-1$ equations (14) and dividing the left sides of the equations by μ_1, we get

$$\left.\begin{array}{l} \beta_{21}\chi_1+\left(\beta_2-\beta_1+\dfrac{\omega_2^2-\omega_1^2}{i\omega_1}\right)\chi_2+\dots+\beta_{2n}\chi_n=0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \beta_{n1}\chi_1+\beta_{n2}\chi_2+\dots+\left(\beta_n-\beta_1+\dfrac{\omega_n^2-\omega_1^2}{i\omega_1}\right)\chi_n=0 \end{array}\right\} \qquad (18)$$

From this system we determine the ratio $\frac{\chi_k}{\chi_1}$ $(k=2, \dots, n)$, discarding second and higher-order terms in β_i, β_{ik}:

$$\frac{\chi_k}{\chi_1}=i\varepsilon_k, \quad \varepsilon_k=-\frac{\beta_{k1}\omega_1}{\omega_k^2-\omega_1^2} \qquad (k=2, \dots, n) \qquad (19)$$

Formulas (19) show that ε_k are small real quantities $(k=2, \dots, n)$. To the root $\mu_1=-\frac{\beta_1}{2}+i\omega_1$ there corresponds the "complex oscillation" (we assume $\chi_1=Ae^{i\alpha}$, $A>0$):

$$\theta_1=\chi_1 e^{\mu_1 t}=Ae^{-\frac{\beta_1}{2}t}e^{i(\omega_1 t+\alpha)},$$

$$\theta_k=\varepsilon_k Ae^{-\frac{\beta_1}{2}t}e^{i\left(\omega_1 t+\alpha+\frac{\pi}{2}\right)} \qquad (k=2, \dots, n) \qquad (20)$$

Taking the linear combination of the given oscillation and the complex conjugate oscillation, we get a "principal" oscillation which corresponds to the roots

$$\mu_1 = -\frac{\beta_1}{2} + i\omega_1 \text{ and } \bar{\mu}_1 = -\frac{\beta_1}{2} - i\omega_1;$$

$$\theta_1 = Ae^{-\frac{\beta_1}{2}t} \sin(\omega_1 t + \alpha),$$

$$\theta_k = \varepsilon_k A e^{-\frac{\beta_1}{2}t} \sin\left(\omega_1 t + \alpha + \frac{\pi}{2}\right) \quad (k = 2, \ldots, n) \tag{21}$$

We get similar expressions for the other principal oscillations as well.

Thus, to a first approximation: (1) small dissipative forces do not alter the frequencies of a conservative system; (2) under these forces the oscillations decay as $t \to \infty$; (3) in the jth principal oscillation all the coordinates are small compared with the jth coordinate and differ from it in phase by a quarter of a period ($j = 1, \ldots, n$).

47. The Effect of a Time-Dependent External Force on Small Oscillations of a Scleronomic System. The Amplitude-Phase Characteristic

In addition to the forces spoken of in Sec. 45, let the forces $Q_i(t)$ ($i = 1, \ldots, n$) act on a scleronomic system. Then the Lagrange equations for small oscillations of the system will differ from the equations (7) on page 229 solely in the presence of nonzero right-hand sides:

$$\sum_{k=1}^{n} (a_{ik}\ddot{q}_k + b_{ik}\dot{q}_k + c_{ik}q_k) = Q_i(t) \quad (i = 1, \ldots, n) \tag{1}$$

The general solution of this nonhomogeneous system of differential equations is given in the form

$$\boldsymbol{q} = \sum_{h=1}^{2n} C_h \boldsymbol{u}_h e^{\mu_h t} + \boldsymbol{q}^* \tag{2}$$

where the first summand is the general solution of the corresponding homogeneous system, and $\boldsymbol{q}^*$ is some particular solution of the system (1).

We assume that the position of the system $q_1 = \ldots = q_n = 0$ is an asymptotically stable position of equilibrium, that is, that $\mathrm{Re}\,\mu_h < 0$ ($h = 1, \ldots, 2n$). Then the first summand tends to zero as t tends to infinity*, and for sufficiently large t the general solution q of the nonhomogeneous system practically coincides with q^*. For this reason, in the future we shall be interested only in "forced oscillations" q^*, which we shall denote simply by q.

Since the system of differential equations (1) is linear, the general case for seeking forced oscillations reduces (at the expense of a linear superposition of particular solutions) to the case when only one of the generalized forces $Q_i(t)$ differs from zero.

For instance, let $Q_1(t) \neq 0$ and $Q_j(t) = 0$ $(j = 2, \ldots, n)$. Also let us first assume that $Q_1(t)$ is a harmonic force, i.e.

$$Q_1(t) = Ae^{i\Omega t} \tag{3}$$

Then the differential equations (1) will be written as

$$\left.\begin{aligned} \sum_{k=1}^{n} (a_{1k}\ddot{q}_k + b_{1k}\dot{q}_k + c_{1k}q_k) &= Ae^{i\Omega t}, \\ \sum_{k=1}^{n} (a_{jk}\ddot{q}_k + b_{jk}\dot{q}_k + c_{jk}q_k) &= 0 \qquad (j = 2, \ldots, n) \end{aligned}\right\} \tag{4}$$

We seek the forced oscillations in the form

$$q_k = B_k e^{i\Omega t} \qquad (k = 1, \ldots, n) \tag{5}$$

Substituting these expressions for q_k $(k = 1, \ldots, n)$ into the differential equations (4) and cancelling $e^{i\Omega t}$, we get the following system of algebraic equations for a determination of the quantities B_k:

$$\left.\begin{aligned} \sum_{k=1}^{n} [a_{1k}(i\Omega)^2 + b_{1k}(i\Omega) + c_{1k}] B_k &= A, \\ \sum_{k=1}^{n} [a_{jk}(i\Omega)^2 + b_{jk}(i\Omega) + c_{jk}] B_k &= 0 \qquad (j = 2, \ldots, n) \end{aligned}\right\} \tag{6}$$

Solving this system, we find for $k = 1, \ldots, n$:

$$B_k = W_{1k}(i\Omega) A \tag{7}$$

where

$$W_{1k}(i\Omega) = \frac{\Delta_{1k}(i\Omega)}{\Delta(i\Omega)} \tag{8}$$

is a regular fractional-rational function of $i\Omega$ with real coefficients; the hodograph of this function in the complex plane, and sometimes the function itself, is termed the *frequency characteristic* or the *amplitude-phase characteristic*.

---

* If the secular equation has multiple roots, then the sum on the right side of (2) can involve secular terms of the form $C_h(u_h + u'_h t + u''_h t^2 + \ldots) e^{\mu_h t}$. However, in this case too the sum tends to zero as t tends to infinity if all $\operatorname{Re} \mu_h < 0$.

Then the response of the coordinate q_k to the external action $Q_1 = Ae^{i\Omega t}$ is obtained by multiplying this action by the frequency characteristic $W_{1k}(i\Omega)$:

$$q_k = W_{1k}(i\Omega)\, Ae^{i\Omega t} \tag{9}$$

Putting

$$W_{1k}(i\Omega) = R_{1k}(\Omega)\, e^{i\Psi_{1k}(\Omega)}$$
$$[R_{1k}(\Omega) > 0] \tag{10}$$

$R_{1k}(\Omega)$ is the amplitude characteristic and $\Psi_{1k}(\Omega)$ is the phase characteristic], we write the formula (9) as follows:

$$q_k = R_{1k}(\Omega)\, Ae^{i[\Omega t + \Psi_{1k}(\Omega)]}$$
$$(k = 1, \ldots, n) \tag{11}$$

Now let

$$Q_1 = A \sin \Omega t \tag{12}$$

i.e.

$$Q_1 = \frac{A}{2i}\left(e^{i\Omega t} - e^{-i\Omega t}\right)$$

The corresponding response will be*

$$q_k = \frac{1}{2i} R_{1k}(\Omega)\, A\left\{e^{i[\Omega t + \Psi_{1k}(\Omega)]} - e^{-i[\Omega t + \Psi_{1k}(\Omega)]}\right\}$$

that is.

$$q_k = R_{1k}(\Omega)\, A \sin[\Omega t + \Psi_{1k}(\Omega)] \tag{13}$$

In other words, when passing from a sinusoidal force (12) to the corresponding response, i.e. to the sinusoidal forced oscillation (13), the amplitude of the force is multiplied by the amplitude characteristic $R_{1k}(\Omega)$, and the phase shift is determined by the phase characteristic $\Psi_{1k}(\Omega)$ taken for the same value of Ω.

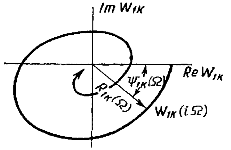

Fig. 53

Fig. 53 depicts the amplitude-phase characteristic $W_{1k}(i\Omega)$ $(0 < \Omega < \infty)$. If for a given Ω the corresponding $R_{1k}(\Omega)$ is very small, then the amplitude of the response is extremely small compared with the amplitude of the "excitation" $A \sin \Omega t$ of the given frequency Ω. On the contrary, if for a given Ω the corresponding $R_{1k}(\Omega)$ is very great, then the response amplitude is great in comparison with the amplitude of the generalized force Q_1. Thus, selecting a system with suitable amplitude characteristics, we can dampen the oscillations in certain frequency ranges and increase the amplitudes of such oscillations on other frequencies. This is the principle on which *filters* are based.

---

* Since $W_{1k}(i\Omega)$ and $W_{1k}(-i\Omega)$ are complex conjugate numbers, it follows that $R_{1k}(-\Omega) = R_{1k}(\Omega)$ and $\Psi_{1k}(-\Omega) = -\Psi_{1k}(\Omega)$.

Since $W(i\Omega)$ is a regular fractional-rational function and for this reason $W(i\Omega) \to 0$ as $\Omega \to \infty$, any system actually passes only a finite range of frequencies.

Now let $Q_1(t)$ be an arbitrary periodic function specified by the Fourier series

$$Q_1(t) = \sum_{m=0}^{\infty} A_m \sin(m\Omega t + \varphi_m) \tag{14}$$

Combining responses to separate harmonics of this series, we get

$$q_k = \sum_{m=0}^{\infty} R_{1k}(m\Omega) A_m \sin[m\Omega t + \varphi_m + \Psi_{1k}(m\Omega)]$$
$$(k = 1, \ldots, n) \tag{15}$$

Now let $Q_1(t)$ be an arbitrary nonperiodic function $f(t)$ of t, which may be given in the form of a Fourier integral

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\Omega t} \, d\Omega \int_{-\infty}^{+\infty} f(t) e^{-i\Omega t} \, dt \tag{16}$$

Putting

$$F(\Omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\Omega t} \, dt \tag{17}$$

we will have

$$Q_1 = f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\Omega) e^{i\Omega t} \, d\Omega \tag{18}$$

The function $F(\Omega)$ is called the *complex spectrum* of the function $Q_1 = f(t)$.

The action

$$\frac{1}{2\pi} F(\Omega) \, d\Omega e^{i\Omega t}$$

will give rise to the response

$$\frac{1}{2\pi} W_{1k}(i\Omega) F(\Omega) \, d\Omega e^{i\Omega t}$$

Therefore, proceeding from the principle of the linear superposition of responses, we find

$$q_k = \frac{1}{2\pi} \int_{-\infty}^{+\infty} W_{1k}(i\Omega) F(\Omega) e^{i\Omega t} \, d\Omega \qquad (k = 1, \ldots, n) \tag{19}$$

that is, the complex spectrum $W_{1k}.(i\Omega)$ $F(\Omega)$ for the coordinate q_k is obtained by multiplying the complex spectrum of the action $Q_1(t)$ by the corresponding frequency characteristic of the system $W_{1k}(i\Omega)$.

Let the system be at rest for $t < 0$, and the motion of the system, given zero initial conditions, be caused solely by the external action $Q(t) \not\equiv 0$ when $t > 0$. Giving the complex spectrum of the response in the form

$$F(\Omega) W_{1k}(i\Omega) = G(\Omega) + iH(\Omega) \tag{20}$$

we will have, according to the formula (19),

$$q_k = \frac{1}{2\pi} \int_{-\infty}^{+\infty} [G(\Omega) + iH(\Omega)] [\cos \Omega t + i \sin \Omega t] \, d\Omega =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} [G(\Omega) \cos \Omega t - H(\Omega) \sin \Omega t] \, d\Omega +$$

$$+ \frac{1}{2\pi} \int_{-\infty}^{+\infty} [G(\Omega) \sin \Omega t + H(\Omega) \cos \Omega t] \, d\Omega$$

Since $q_k(t)$ is a real function, the second summand on the right is zero* and for this reason

$$q_k = \frac{1}{2\pi} \int_{-\infty}^{+\infty} [G(\Omega) \cos \Omega t - H(\Omega) \sin \Omega t] \, d\Omega \tag{21}$$

Now suppose that

$$\left.\begin{array}{ll} Q_1(t) = 0, & \\ q_k(t) = 0 & (k = 1, \ldots, n) \end{array}\right\} \text{for } t < 0 \tag{22}$$

then

$$0 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} [G(\Omega) \cos \Omega t - H(\Omega) \sin \Omega t] \, d\Omega \quad (\text{for } t < 0)$$

---

* A real system cannot have a frequency characteristic that does not satisfy this condition.

Replacing here t by $-t$, we have

$$0=\frac{1}{2\pi}\int_{-\infty}^{+\infty}[G(\Omega)\cos\Omega t+H(\Omega)\sin\Omega t]\,d\Omega\quad(\text{for } t>0)\tag{23}$$

Adding the equations (21) and (23) termwise, we have

$$q_k=\frac{1}{\pi}\int_{-\infty}^{+\infty}G(\Omega)\cos\Omega t\,d\Omega\quad(\text{for } t>0)\tag{24}$$

It may be shown that the expression (24) is the general solution of a stable system for the coordinate q_k, for the conditions (22)

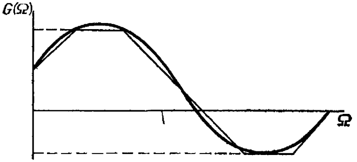

Fig. 54

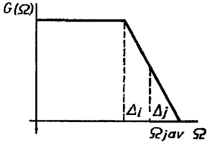

Fig. 55

and for zero initial conditions *. The function $W_{1k}(\Omega)$ is constructed directly from the coefficients a_{ik}, b_{ik} and c_{ik}, which appear in equations (1), and the function $F(\Omega)$ is determined by the expression (17). After that the problem of determining the motion described by equations (1) reduces to a single quadrature with the help of (24).

For this reason, any method of approximate integration is suitable for an approximate determination of the motion. For example, it is possible to construct a graph of the function $G(\Omega)$, replace it with a polygonal line and through the salient points draw horizontal lines to the axis of ordinates. Then the function $G(\Omega)$ is approximately replaced by the sum of the functions $g_j(\Omega)$, the graph of each of which is a trapezoid (a triangle, in the special case), as shown in Fig. 54.

The integral (24) may be computed for one such function $g_j(\Omega)$, as follows:

$$q_{jk}=\frac{1}{\pi}\int_{-\infty}^{\infty}g_j(\Omega)\cos\Omega t\,d\Omega=\frac{A_j}{\pi}\,\frac{\sin\Omega_{jav}t}{\Omega_{jav}}\,\frac{\sin\Delta_j t}{\Delta_j t}$$

---

* This follows from the fact that expression (20) defines the Fourier transform of the solution for zero initial conditions, and expression (24) defines the inverse Fourier transform.

where Ω_{jav} and Δ_j are shown in Fig. 55, and A_j is the area of the trapezoid. Therefore

$$q_k \approx \frac{1}{\pi} \sum_j A_j \frac{\sin \Omega_{jav} t}{\Omega_{jav}} \frac{\sin \Delta_j t}{\Delta_j t}$$

where the summation is taken over all trapezoids obtained in the approximation of $G(\Omega)$ by a polygonal line.

Tables of the functions $\frac{\sin z}{z}$ may be used for such an approximation when constructing motions.

CHAPTER 7

Systems with Cyclic Coordinates

48. Reduced System. The Routh Potential. Hidden Motions. Hertz' Conception of the Kinetic Origin of Potential Energy

In the present chapter the general propositions given in Chapters 2 and 5 are used for investigating the motion of a holonomic scleronomic system with cyclic coordinates q_α $(\alpha = m + 1, \ldots, n)$. The kinetic energy of such a system is of the form

$$T = \frac{1}{2} \sum_{r,\, s=1}^{n} a_{rs}(q_1, \ldots, q_m)\, \dot{q}_r \dot{q}_s \tag{1}$$

Let us find an expression for the kinetic energy in Routh variables $q_i, \dot{q}_i, p_\alpha$ $(i = 1, \ldots, m;\ \alpha = m + 1, \ldots, n)$. To do this, express all $\dot{q}_\alpha$ in terms of p_α $(\alpha = m + 1, \ldots, n)$, utilizing the initial relations

$$p_\alpha = \frac{\partial T}{\partial \dot{q}_\alpha} = \sum_{i=1}^{m} a_{\alpha i} \dot{q}_i + \sum_{\beta=m+1}^{n} a_{\alpha\beta} \dot{q}_\beta \quad (\alpha = m + 1, \ldots, n) \tag{2}$$

Since the determinant $D = \det (a_{\alpha\beta})^{n}_{\alpha,\, \beta=m+1} \neq 0)$*, from relations (2) we get

$$\dot{q}_\alpha = \sum_{\beta=m+1}^{n} b_{\alpha\beta} \left(p_\beta - \sum_{i=1}^{m} a_{\beta i} \dot{q}_i\right) \quad (\alpha = m + 1, \ldots, n) \tag{3}$$

where $\| b_{\alpha\beta} \|^{n}_{m+1}$ is the inverse matrix of the matrix $\| a_{\alpha\beta} \|^{n}_{m+1}$,

$$\| b_{\alpha\beta} \| = \| a_{\alpha\beta} \|^{-1} \tag{4}$$

Setting

$$\gamma_{\alpha i} = \sum_{\beta=m+1}^{n} b_{\alpha\beta} a_{\beta i} \quad (i = 1, \ldots, m;\ \alpha = m + 1, \ldots, n) \tag{5}$$

we write (3) in the following form:

$$\dot{q}_\alpha = \sum_{\beta=m+1}^{n} b_{\alpha\beta} p_\beta - \sum_{i=1}^{m} \gamma_{\alpha i} \dot{q}_i \quad (\alpha = m + 1, \ldots, n) \tag{6}$$

---

* See footnote on page 78.

Here, $b_{\alpha\beta}$ and $\gamma_{\alpha i}$ are functions of noncyclic (or, as they are sometimes called, position) coordinates $q_i\ (i=1, \dots, m)$. Putting expressions (6) for $\dot q_\alpha$ into formula (1), we get the expression $\widehat{T}$ for the kinetic energy in Routh variables:

$$\widehat{T}=\frac{1}{2}\sum_{i,\,j=1}^{m} a^*_{ij}\dot q_i\dot q_j+\frac{1}{2}\sum_{\alpha,\,\beta=m+1}^{n} a^*_{\alpha\beta}p_\alpha p_\beta+\sum_{i=1}^{m}\sum_{\alpha=m+1}^{n} a^*_{i\alpha}\dot q_i p_\alpha \qquad (7)$$

A remarkable circumstance (one noticed by Routh) is that in this formula all the $a^*_{i\alpha}=0$, i.e. the expression for $\widehat{T}$ is the sum of the quadratic form in the position velocities $\dot q_1, \dots, \dot q_m$ and the quadratic form in the generalized momenta $p_{m+1}, \dots, p_n$.*

Indeed,

$$a^*_{i\alpha}=\frac{\partial^2\widehat{T}}{\partial\dot q_i\,\partial p_\alpha}=\frac{\partial}{\partial\dot q_i}\left(\sum_{\beta=m+1}^{n}\frac{\partial T}{\partial\dot q_\beta}\frac{\partial\dot q_\beta}{\partial p_\alpha}\right)=\frac{\partial}{\partial\dot q_i}\sum_{\beta=m+1}^{n} b_{\beta\alpha}p_\beta=0 \qquad (8)$$

since $\sum\limits_{\beta=m+1}^{n} b_{\alpha\beta}p_\beta$ is dependent only on the variables q_i and p_β, which are regarded as independent with respect to the $\dot q_i\ (i=1, \dots, m;\ \beta=m+1, \dots, n)$.

Now compute the coefficients $a^*_{\alpha\beta}$:

$$a^*_{\alpha\beta}=\frac{\partial^2\widehat{T}}{\partial p_\alpha\,\partial p_\beta}=\frac{\partial}{\partial p_\alpha}\sum_{\gamma=m+1}^{n}\frac{\partial T}{\partial\dot q_\gamma}\frac{\partial\dot q_\gamma}{\partial p_\beta}=\frac{\partial}{\partial p_\alpha}\sum_{\gamma=m+1}^{n} b_{\gamma\beta}p_\gamma=b_{\alpha\beta}$$
$$(\alpha,\beta=m+1, \dots, n) \qquad (9)$$

Similarly, we find** the coefficients a^*_{ij}

$$a^*_{ij}=\frac{\partial^2\widehat{T}}{\partial\dot q_i\,\partial\dot q_j}=\frac{\partial}{\partial\dot q_i}\left[\frac{\partial T}{\partial\dot q_j}+\sum_{\alpha=m+1}^{n}\frac{\partial T}{\partial\dot q_\alpha}\frac{\partial\dot q_\alpha}{\partial\dot q_j}\right]_{\dot q_\alpha=\dots}=$$
$$=\frac{\partial}{\partial\dot q_i}\left[\frac{\partial T}{\partial\dot q_j}\right]_{\dot q_\alpha=\dots}-\frac{\partial}{\partial\dot q_i}\sum_{\alpha=m+1}^{n}\gamma_{\alpha i}p_\alpha=$$

---

[*] The transformation of variables used here that corresponds to a transition from the variables $\dot q_i$, $\dot q_\alpha$ to the variables $\dot q_i$, p_α is ordinarily used in the theory of quadratic forms to reduce (by Lagrange's method) a quadratic form to a sum of squares. Indeed, applying such transformations a number of times, we can represent a quadratic form in n variables in the form of a sum of n quadratic forms, each of which is equal to the product of the square of this variable by some real coefficient.

[**] Here and henceforward $\dot q_\alpha = \dots$ signifies that all the $\dot q_\alpha$ in the expression in square brackets are replaced by their expressions (6).

$$= \frac{\partial}{\partial \dot{q}_i}\left[\frac{\partial T}{\partial \dot{q}_j}\right]_{\dot{q}_\alpha = \dots} = \frac{\partial^2 T}{\partial \dot{q}_i\,\partial \dot{q}_j} + \sum_{\alpha=m+1}^{n} \frac{\partial^2 T}{\partial \dot{q}_j\,\partial \dot{q}_\alpha}\,\frac{\partial \dot{q}_\alpha}{\partial \dot{q}_i} =$$

$$= a_{ij} - \sum_{\alpha=m+1}^{n} \gamma_{\alpha i} a_{\alpha j}$$

Utilizing equations (5) we get

$$a^*_{ij} = a_{ij} - \sum_{\alpha,\,\beta=m+1}^{n} b_{\alpha\beta} a_{\alpha j} a_{\beta j} \quad (i, j = 1, \dots, m) \tag{10}$$

But on the basis of (4), $b_{\alpha\beta} = \frac{A_{\beta\alpha}}{D}$, where $A_{\beta\alpha}$ is the cofactor of the element $a_{\beta\alpha}$ in the determinant D. Taking advantage of this circumstance, in place of (10) we can write

$$a^*_{ij} = \frac{1}{D}\begin{vmatrix} a_{ij} & a_{i,m+1} & \dots & a_{in} \\ a_{m+1,j} & a_{m+1,m+1} & \dots & a_{m+1,n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{nj} & a_{n,m+1} & \dots & a_{nn} \end{vmatrix} \quad (i, j = 1, \dots, m) \tag{10'}$$

Thus, formula (7) has the form

$$\widehat{T} = \frac{1}{2}\sum_{i,j=1}^{m} a^*_{ij}\dot{q}_i\dot{q}_j + \frac{1}{2}\sum_{\alpha,\,\beta=m+1}^{n} b_{\alpha\beta}p_\alpha p_\beta \tag{11}$$

Here, the coefficients a^*_{ij} and $b_{\alpha\beta}$ are determined by the equations (4) and (10).

Let the forces impressed on the scleronomic system have the potential $\Pi = \Pi(t, q_i)$. Then, $L = T - \Pi$. Compute the Routh function (see Sec. 13):

$$R = R(t, q_i, \dot{q}_i, p_\alpha) = \sum_{\alpha=m+1}^{n} p_\alpha \dot{q}_\alpha - L =$$

$$= \sum_{\alpha=m+1}^{n} p_\alpha\left(\sum_{\beta=m+1}^{n} b_{\alpha\beta}p_\beta - \sum_{i=1}^{m}\gamma_{\alpha i}\dot{q}_i\right) - T + \Pi =$$

$$= -\frac{1}{2}\sum_{i\ j=1}^{m} a^*_{ij}\dot{q}_i\dot{q}_j + \frac{1}{2}\sum_{\alpha,\,\beta=m+1}^{n} b_{\alpha\beta}p_\alpha p_\beta + \Pi -$$

$$-\sum_{i=1}^{m}\left(\sum_{\alpha=m+1}^{n}\gamma_{\alpha i}p_\alpha\right)\dot{q}_i \tag{12}$$

Consider the function $\Pi^*(t, q_i, p_\alpha)$, which we shall call the *Routh potential*:*

$$\Pi^* = \Pi + \frac{1}{2} \sum_{\alpha,\, \beta=m+1}^{n} b_{\alpha\beta} p_\alpha p_\beta \tag{13}$$

Taking advantage of the fact that

$$b_{\alpha\beta} = \frac{A_{\beta\alpha}}{D}$$

where $A_{\beta\alpha}$ is the cofactor of the element $a_{\beta\alpha}$ in the determinant D, we can write the expression for the Routh potential also in the following form:

$$\Pi^* = \Pi - \frac{1}{2D} \begin{vmatrix} 0 & p_{m+1} & \dots & p_n \\ p_{m+1} & a_{m+1,\, m+1} & \dots & a_{m+1,\, n} \\ \vdots & \vdots & & \vdots \\ p_n & a_{n,\, m+1} & \dots & a_{nn} \end{vmatrix} \tag{13'}$$

We also introduce the notation

$$T^* = \frac{1}{2} \sum_{i,\, j=1}^{m} a_{ij}^* \dot{q}_i \dot{q}_j \tag{14}$$

Then, according to (12), the function $-R$, which for the position coordinates plays the role of the Lagrangian function, is

$$T^* - V \tag{15}$$

where V is the generalized potential defined by the equation

$$V = -\sum_{i=1}^{m} \Big(\sum_{\alpha=m+1}^{n} \gamma_{\alpha i} p_\alpha \Big) \dot{q}_i + \Pi^*(t, q_i, p_\alpha) \tag{16}$$

Let us now consider some motion of the initial system. In this motion,

$$p_\alpha = \mathrm{const} = c_\alpha \qquad (\alpha = m+1, \dots, n) \tag{17}$$

and the variation of the position coordinates $q_i = q_i(t)$ $(i = 1, \dots \dots, m)$ may be determined from the differential equations (8) on page 81, in which p_α should everywhere be replaced by c_α $(\alpha = m + {} + 1, \dots, n)$. But these equations are Lagrange equations (with the Lagrangian function $-R = T^* - V$) for a certain auxiliary natural scleronomic system with m degrees of freedom having the kinetic

---

[*] Routh E. G., *The Advanced Part of a Treatise on the Dynamics of a System of Rigid Bodies*, 5th ed., 1892, Sec. 99.

energy $T^* = \frac{1}{2} \sum_{i,j=1}^{m} a_{ij}^* \dot{q}_i \dot{q}_j$ and the generalized force potential $V = -\sum_{i=1}^{m} \left(\sum_{\alpha=m+1}^{n} \gamma_{\alpha i} c_\alpha \right) \dot{q}_i + \Pi^*(t, q_i, c_\alpha)$. The Routh potential $\Pi^*(t, q_i, c_\alpha)$ is the potential energy of this system. From formulas (11) to (13) it follows that

$$T + \Pi = T^* + \Pi^* \tag{18}$$

We shall call this auxiliary system a *reduced system.*

Thus, *the variation of the position coordinates $q_1, \ldots, q_m$ determines the motion of a reduced system (with m degrees of freedom) with kinetic energy T^* and with generalized potential $V = V_1 + \Pi^*$, where Π^* is the Routh potential. Given appropriate motions of the initial and reduced systems, the total energies of these systems are equal.*

When the functions $q_i = q_i(t)$ $(i = 1, \ldots, m)$ which define the motion of a reduced system are found, then the time variation of the cyclic coordinates $q_\alpha = q_\alpha(t)$ $(\alpha = m + 1, \ldots, n)$ may be determined from the formula (9) on page 81, which can now be written as

$$q_\alpha = \int \left[\frac{\partial \Pi^*(t, q_i, c_\alpha)}{\partial c_\alpha} - \sum_{j=1}^{m} \gamma_{\alpha j}(q_i) \dot{q}_j \right] dt + c_\alpha \tag{19}$$

$$(\alpha = m + 1, \ldots, n)$$

Consider a special case when the expression for the kinetic energy of the initial system (1) does not contain products of the position velocities $\dot{q}_i$ by the cyclic velocities $\dot{q}_\alpha$, i.e. a case when all $a_{i\alpha} = 0$ $(i = 1, \ldots, m;\ \alpha = m + 1, \ldots, n)$. In this case, the kinetic energy T breaks up into two quadratic forms, of which one contains only the position velocities $\dot{q}_i$, while the other contains only the cyclic velocities $\dot{q}_\alpha$. Such a system is termed *gyroscopically uncoupled.* For a gyroscopically uncoupled system, according to formulas (5), all the $\gamma_{\alpha i} = 0$, and hence, $V_1 = 0$ and $V = \Pi^*$. Thus, *if the initial system is gyroscopically uncoupled, the reduced system has the ordinary potential* Π^*. *

---

* The reduced system is acted on by potential forces $-\frac{\partial \Pi^*}{\partial q_i}$ and gyroscopic forces determined by the potential V_1. In the case of a gyroscopically free system $V_1 = 0$ and gyroscopic forces are absent.

Besides, from (10) it follows that for the gyroscopically uncoupled system $a_{ij}^* = a_{ij}$ $(i, j = 1, \ldots, m)$, i.e.

$$T^* = \frac{1}{2} \sum_{i, j=1}^{m} a_{ij} \dot{q}_i \dot{q}_j \tag{20}$$

In this case, the initial system has the kinetic energy

$$T = \frac{1}{2} \sum_{i, j=1}^{m} a_{ij} \dot{q}_i \dot{q}_j + \frac{1}{2} \sum_{\alpha, \beta=m+1}^{n} b_{\alpha\beta} c_\alpha c_\beta$$

and the potential energy $\Pi = \Pi(t, q_i)$ while the reduced system has the kinetic energy $T^* = \frac{1}{2} \sum\limits_{i, j=m+1}^{n} a_{ij} \dot{q}_i \dot{q}_j$ and potential energy

$$\Pi^* = \Pi(t, q_i) + \frac{1}{2} \sum_{\alpha, \beta=m+1}^{n} b_{\alpha\beta} c_\alpha c_\beta$$

We see that *part of the kinetic energy of the initial system* $\left(\frac{1}{2} \sum\limits_{\alpha, \beta=m+1}^{n} b_{\alpha\beta} c_\alpha c_\beta\right)$ *has passed into the potential energy of the reduced system.*

All that has been said about scleronomic systems holds true, in particular, for conservative systems in which $\frac{\partial \Pi}{\partial t} = 0$ i. e. $\Pi = \Pi(q_i)$. The reduced system for a gyroscopic uncoupled conservative system will again be conservative.

Now let there be given an arbitrary conservative system with m degrees of freedom, with kinetic energy $T^* = \frac{1}{2} \sum\limits_{i, j=1}^{m} a_{ij} \times (q_1, \ldots, q_m) \dot{q}_i \dot{q}_j$ and potential energy $\Pi^* = \Pi^*(q_i)$. Consider a conservative system in which the number of degrees of freedom is $m + 1$ and which has m position coordinates $q_1, \ldots, q_m$ and one cyclic coordinate q_{m+1}.

In the new system let $\Pi = 0$ and the kinetic energy have the form

$$T = T^* + \frac{1}{2} \frac{k}{\Pi^*(q_i)} \dot{q}_{m+1}^2 \qquad (k = \text{const})$$

Then

$$p_{m+1} = \frac{k}{\Pi^*(q_i)} \dot{q}_{m+1}$$

and

$$T = T^* + \frac{1}{2k} \Pi^*(q_i) p_{m+1}^2$$

But then for $p_{m+1} = c_{m+1} = \sqrt{2k}$

$$\frac{1}{2}\Pi^*(q_i)\,p^2_{m+1} = \Pi^*(q_i)$$

Thus, in the given conservative system, T^* is the kinetic, and $\Pi^* = \Pi^*(q_i)$ is the potential energy, whereas in the new "extended" system we have, respectively, $T = T^* + \Pi^*(q_i)$ and $\Pi = 0$. *The potential energy of the given system results from the kinetic energy of the "extended" system which has a greater number of degrees of freedom.*

Motions, under which only cyclic (hidden) coordinates vary are called hidden motions. *

Above we saw that the potential energy of a conservative system can always be regarded as the kinetic energy of hidden motions.

This conception about the kinetic origin of potential energy and, hence, about the kinetic origin of forces applied to bodies performing explicit (nonhidden) motions was broadly developed by Hertz in 1894. **

Example. Let us consider the motion of a rigid body with fixed point O in the Lagrangian case when the body is acted upon by the force of gravity Mg, when there is an axis of dynamic symmetry, and the centre of gravity D is located on this axis.

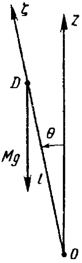

Fig. 56

We shall specify the rigid body by means of the three Euler angles θ, ψ, φ, where the angle of nutation θ is the angle between the vertical axis Oz and the axis of dynamic symmetry $O\zeta$ (Fig. 56), ψ is the angle of precession, and φ is the angle of pure rotation. Let $l = OD$ and A and C be the equatorial and axial moments of inertia, respectively. Then

---

* Systems containing within them gyroscopes are a brilliant example of the manifestation of hidden motions. Because of these "hidden" motions of gyroscopes, the motion of such systems differs radically from the motion of systems without gyroscopes.

** H. Hertz, *Die Prinzipien der Mechanik in neuem Zusammenhänge dargestellt*, see also W. Thomson and P. Tait, *Treatise on Natural Philosophy*, Part 1, Sec. 319 and A. G. Webster, *The Dynamics of Particles and of Rigid, Elastic and Fluid Bodies*, 1925.

$$2T = A(p^2+q^2)+Cr^2, \quad \Pi = Mgl\cos\theta,$$

where p, q, r are the projections of the angular velocity ω on the principal axes of inertia $O\xi$, $O\eta$, $O\zeta$. But

$$p^2+q^2=\dot{\theta}^2+\dot{\psi}^2\sin^2\theta, \quad r=\dot{\varphi}+\dot{\psi}\cos\theta;$$

therefore, we finally have

$$\begin{gathered}2T=A(\dot{\theta}^2+\dot{\psi}^2\sin^2\theta)+C(\dot{\varphi}+\dot{\psi}\cos\theta)^2,\\ \Pi=Mgl\cos\theta\end{gathered} \tag{21}$$

Find expressions for the generalized momenta p_φ, p_ψ corresponding to the cyclic coordinates φ, ψ:

$$\left.\begin{aligned}p_\psi&=(A\sin^2\theta+C\cos^2\theta)\dot{\psi}+C\cos\theta\dot{\varphi},\\ p_\varphi&=C\cos\theta\dot{\psi}+C\dot{\varphi}\end{aligned}\right\} \tag{22}$$

Solving these relations for $\dot{\psi}$ and $\dot{\varphi}$, we have

$$\left.\begin{aligned}\dot{\psi}&=\frac{p_\psi-\cos\theta p_\varphi}{A\sin^2\theta},\\ \dot{\varphi}&=\frac{-C\cos\theta p_\psi+(A\sin^2\theta+C\cos^2\theta)p_\varphi}{AC\sin^2\theta}\end{aligned}\right\} \tag{23}$$

The coefficients of p_ψ and p_φ on the right sides of (23) are the quantities $b_{\alpha\beta}$. The expression for the kinetic energy in the Routh variables θ, $\dot{\theta}$, p_ψ, p_φ is obtained by substituting expressions (23) for $\dot{\varphi}$ and $\dot{\psi}$ into (21) for $2T$ or straightway by formula (11)*:

$$2T=A\dot{\theta}^2+\frac{Cp_\psi^2-2C\cos\theta p_\psi p_\varphi+(A\sin^2\theta+C\cos^2\theta)p_\varphi^2}{AC\sin^2\theta} \tag{24}$$

When the system is in motion, the momenta p_ψ and p_φ retain constant values:

$$p_\psi=a, \quad p_\varphi=b \tag{25}$$

The nutational motion is determined by the reduced system, for which

$$T^*=\frac{1}{2}A\dot{\theta}^2, \quad \Pi^*=Mgl\cos\theta+\frac{(a-b\cos\theta)^2}{A\sin^2\theta}+\frac{b^2}{C} \tag{26}$$

The change in the angle of nutation $\theta=\theta(t)$ is found from the corresponding energy integral

$$\frac{1}{2}A\dot{\theta}^2+Mgl\cos\theta+\frac{(a-b\cos\theta)^2}{A\sin^2\theta}=\text{const}=h \tag{27}$$

---

* In the given case the rigid body is a gyroscopically uncoupled system, so that in formula (11) $a_{ij}^*=a_{ij}$ $(i,\ j=1,\ \ldots,\ m)$.

Introduce the auxiliary variable $u = \cos\theta$. Multiplying both sides of (27) by $\frac{2}{A}(1-u^2) = \frac{2}{A}\sin^2\theta$, we find

$$\left(\frac{du}{dt}\right)^2 = f(u), \tag{28}$$

where $f(u)$ is a third-degree polynomial in u:

$$f(u) = \frac{2}{A}(h - Mglu)(1-u^2) - \frac{2}{A^2}(a-bu)^2 \tag{29}$$

Putting $u = \pm 1$ and $u = u_0 = \cos\theta_0 < 1$, we find*

$$f(-1) < 0,\ f(u_0) = \left(\frac{du}{dt}\right)_0^2 > 0,\ f(+1) < 0$$

Then the polynomial $f(u)$ has three real roots $u_1 = \cos\theta_1$, $u_2 = \cos\theta_2$ and u':

$$-1 < u_1 < u_2 < 1 < u'$$

and [since $f(+\infty) = +\infty$] the graph of the polynomial $f(u)$ has the form shown in Fig. 57.

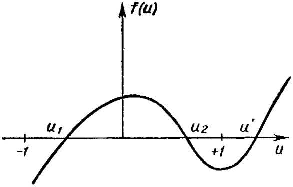

Fig. 57

Since when the body is in motion, $-1 \leqslant u = \cos\theta \leqslant +1$ and $f(u) = \left(\frac{du}{dt}\right)^2 \geqslant 0$, it follows that $u = \cos\theta$ must vary in the interval $u_1 \leqslant u \leqslant u_2$, i.e. $\theta_1 \geqslant \theta \geqslant \theta_2$. The angular velocity of precession is determined from the first formula in (23):

$$\dot{\psi} = \frac{a - b\cos\theta}{A\sin^2\theta} \tag{30}$$

If the ratio $\frac{a}{b}$ lies outside the interval $u_1 \leqslant u \leqslant u_2$, then the velocity of precession $\dot{\psi}$ retains a constant sign and precessional motion occurs all the time in one direction. In this case the trace left by the axis $O\zeta$ on a fixed sphere with centre at O describes a curve depicted in Fig. 58.

---

* Here we assume that $a \neq \pm b$. The special cases $a = \pm b$ will be considered below. Besides, the initial value θ_0 is chosen so that $\dot{\theta}_0 \neq 0$ and, hence, $\left(\frac{du}{dt}\right)_0 \neq 0$.

But if $u_1 < \frac{a}{b} < u_2$, then the velocity of precession changes sign when $\cos\theta^* = \frac{a}{b}$, and the trace of the dynamic axis describes on the sphere a curve as shown in Fig. 59.

If $\frac{a}{b} = u_2$, then the velocity of precession $\dot{\psi}$ does not change sign, but vanishes on the upper parallel $\theta = \theta_2$ (Fig. 60).*

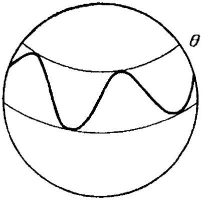

Fig. 58

Analytically, the change in angle of nutation is found from the formulas

$$\pm \int \frac{du}{\sqrt{f(u)}} = t + \text{const}, \quad \theta = \arccos u \tag{31}$$

The plus sign is taken when θ varies from θ_2 to θ_1 and the minus sign when it varies in the reverse direction. It is obvious that the variation in time of the angle θ will be a periodic function $\theta\ (t)$ with period

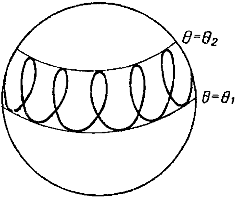

Fig. 59

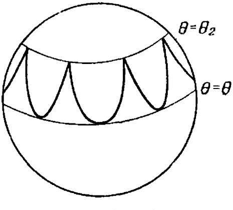

Fig. 60

$$\tau = 2 \int\limits_{u_1}^{u_2} \frac{du}{\sqrt{f(u)}} \tag{32}$$

After the variation of the angle of nutation $\theta = \theta\ (t)$ is found, the variations of the angles ψ and φ are determined from the formulas (23).

---

* It may be proved that $\dot{\psi}$ cannot vanish on the lower parallel.

Let us now consider the motion of a rigid body passing through the "singular" position $\theta = 0$. At this point the kinetic energy is given by the degenerate quadratic form [see expression (21) for $\theta = 0$]

$$2T = A\dot{\theta}^2 + C(\dot{\psi} + \dot{\varphi})^2 \tag{33}$$

According to formulas (22), the equation $p_\psi = p_\varphi$ that is

$$a = b \tag{34}$$

should occur at the singular point $\theta = 0$. Assuming that the arbitrary constants a and b are connected by the relation (34), we readily evaluate the indeterminate form in the expression (26) for the Routh potential:*

$$\Pi^* = Mgl\cos\theta + \frac{a^2}{A}\,\frac{1-\cos\theta}{1+\cos\theta}$$

Then the nutational motion is again determined from the energy integral $T^* + \Pi^* = \text{const}$.

If the motion passes through the "singular" position $\theta = \pi$, then in place of (34) we will have the relation

$$a = -b$$

and the expression for the Routh potential will take the form

$$\Pi^* = Mgl\cos\theta + \frac{a^2}{A}\,\frac{1+\cos\theta}{1-\cos\theta}$$

In the above considered two "singular" cases, the upper (or, respectively, lower) parallel in Figs. 58-60 degenerates into a point.

49. Stability of Stationary Motions

Consider a conservative system, the position of which is specified by m position coordinates q_i and $n - m$ cyclic coordinates q_α ($i = 1, \ldots, m$; $\alpha = m + 1, \ldots, n$). The motion of such a system is defined by the canonical equations

$$\frac{dq_r}{dt} = \frac{\partial H}{\partial p_r}, \quad \frac{dp_r}{dt} = -\frac{\partial H}{\partial q_r} \quad (r = 1, \ldots, n), \tag{1}$$

where the total energy of the system $H = H(q_i, p_i, p_\alpha)$ has the form

$$H = \frac{1}{2}\sum_{r,s=1}^{n} c_{rs}(q_1, \ldots, q_m)\,p_r p_s + \Pi(q_1, \ldots, q_m) \tag{2}$$

where the quadratic form on the right side is positive definite **. When the system is in motion, the function H and the generalized momenta p_α ($\alpha = m + 1, \ldots, n$) do not alter their values, i.e.

---

* In the expression for Π^* we discard the constant term $\frac{b^2}{B}$.

** This quadratic form is the kinetic energy of the system expressed in terms of generalized momenta.

these quantities are the integrals of motion:

$$\left.\begin{aligned} &H(q_i, p_i, p_\alpha) = H(q_i^0, p_i^0, p_\alpha^0) \\ &p_\alpha = p_\alpha^0 \quad (\alpha = m+1, \ldots, n) \end{aligned}\right\} \tag{3}$$

The motion of a system is called *stationary* if during it all the position coordinates retain constant values $q_i = \text{const} = q_i^0$ $(i = 1, \ldots, m)$.

In stationary motion all position velocities are zero and therefore, by equations (1) and equality (2),

$$\dot{q}_i = \sum_{k=1}^{m} c_{ik}(q_i^0)\, p_k + \sum_{\alpha=m+1}^{n} c_{i\alpha}(q_i^0)\, p_\alpha^0 = 0 \quad (i = 1, \ldots, m) \tag{4}$$

From these relations it follows,* that in stationary motion all position momenta are also of constant magnitude: $p_i = \text{const} = p_i^0$. Since $\dot{q}_i = \dot{p}_i = 0$ $(i = 1, \ldots, m)$, it follows from (1) that for stationary motion

$$\frac{\partial H}{\partial q_i} = 0, \quad \frac{\partial H}{\partial p_i} = 0 \quad (i = 1, \ldots, m) \tag{5}$$

Equations (5) are necessary and sufficient conditions which the initial values q_i^0, p_i^0, p_α^0 $(i = 1, \ldots, m;\ \alpha = m+1, \ldots, n)$ must satisfy for the motion defined by them to be stationary.

Note also that in stationary motion the cyclic velocities $\dot{q}_\alpha$ $(\alpha = m+1, \ldots, n)$ also have constant magnitude, since by (1)

$$\dot{q}_\alpha = \frac{\partial H}{\partial p_\alpha} = \sum_{s=1}^{n} c_{\alpha s}(q^0)\, p_s^0 = \text{const} \quad (\alpha = m+1, \ldots, n) \tag{6}$$

Therefore

$$q_\alpha = \dot{q}_\alpha^0 (t - t_0) + q_\alpha^0 \quad (\alpha = m+1, \ldots, n) \tag{7}$$

The initial cyclic coordinates q_α^0 are not arbitrary constants and do not enter into the conditions (5).

Thus, *in stationary motion the position coordinates retain constant value while the cyclic coordinates vary in accordance with a linear law.*

Example. The regular precession of a heavy symmetric gyroscope is an instance of stationary motion.

Indeed, regular precession is characterized by the equations

$$\theta = \text{const} = \theta_0, \quad \dot{\psi} = \text{const}, \quad \dot{\varphi} = \text{const}$$

where the angle of precession ψ and the angle of pure rotation φ are cyclic coordinates, while the angle of nutation θ—an angle formed by the axis of the gyroscope and the vertical—is a position coordinate.*

---

* Relations (4) may be solved for the position momenta p_k, since the determinant $\det\,[c_{ik}(q_j^0)]_{i,k=1}^{m}$ is different from zero, inasmuch as the quadratic form $\sum_{r,s=1}^{n} c_{rs} p_r p_s$ is positive definite.

Note that in accordance with relations (6), a slight change in the initial quantities q_i^0, p_i^0, p_α^0 yields a small change in the original cyclic velocities $\dot{q}_\alpha^0$ $(\alpha = m + 1, \ldots, n)$. However, a small change in the quantities $\dot{q}_\alpha^0$, according to formula (7), yields, as time progresses, an arbitrarily large change in the cyclic coordinates themselves. For this reason, *stationary motion cannot be stable with respect to cyclic coordinates.*

Henceforward, by stability of stationary motion we shall mean the stability with respect to all momenta p_i and p_α and position coordinates q_i $(i = 1, \ldots, m;\ \alpha = m + 1, \ldots, n)$.**

We then have the following **criterion of stability of stationary motion.**

Motion with the initial conditions q_i^0, p_i^0, p_α^0 $(i = 1, \ldots, m;$ $\alpha = m + 1, \ldots, n)$ *is stable stationary motion if the function* $H(q_i, p_i, p_\alpha^0)$ *at the point* $q_i = q_i^0$, $p_i = p_i^0$ $(i = 1, \ldots, m)$ *has a strict extremum.*

Indeed, the motion under consideration will be stationary inasmuch as it follows from the existence of the extremum of the function $H(q_i, p_i, p_\alpha^0)$ that the quantities q_i^0, p_i^0 $(i = 1, \ldots, m)$ together with the quantities p_α^0 $(\alpha = m + 1, \ldots, n)$ satisfy equations (5). We introduce the deviations

$$\xi_i = q_i - q_i^0, \quad \eta_i = p_i - p_i^0, \quad \eta_\alpha = p_\alpha - p_\alpha^0.$$

Then, utilizing the motion integral $H(q_i^0 + \xi_i, p_i^0 + \eta_i, p_\alpha^0)$ as the Lyapunov function, we can conclude (see page 185) that the zero solution $\xi_i = 0$, $\eta_i = 0$ (i.e. the stationary motion) is stable on the assumption that the cyclic momenta p_α do not experience perturbations (i.e. that these quantities have the same values p_α^0 for perturbed motion as for unperturbed motion). ***

However, the above-formulated criterion of stability is valid in the general case as well, when for perturbed motion the quantities $\eta_\alpha = p_\alpha - p_\alpha^0$ $(\alpha = m + 1, \ldots, n)$ are different from zero. To see this, it will suffice to use a corollary to the Lyapunov theorem (page 185) taking the function

$$[H(q_i^0 + \xi_i, p_i^0 + \eta_i, p_\alpha^0) - H(q_i^0, p_i^0, p_\alpha^0)]^2 + \sum_{\alpha=m+1}^{n} \eta_\alpha^2 \qquad (8)$$

---

* From (21) of the preceding section it is seen that the coordinates φ and ψ do not appear explicitly in the expressions for the kinetic energy and the potential energy.

** Or, which is the same thing, the stability with respect to the quantities q_i, $\dot{q}_i$, p_α $(i = 1, \ldots, m;\ \alpha = m + 1, \ldots, n)$.

*** This is precisely the kind of stability that Routh considered.

for the Lyapunov function. This function is the integral of motion and has a strict minimum equal to zero at $\xi_i = 0$, $\eta_i = 0$, $\eta_\alpha = 0$ $(i = 1, \ldots, m;\ \alpha = m + 1, \ldots, n)$.

Let us establish an analogy between the stability of the state of equilibrium and the stability of a stationary motion. Consider a reduced system with m independent coordinates $q_1, \ldots, q_m$ and with potential energy equal to the Routh potential:

$$\Pi^* (q_i, p_\alpha^0) = \Pi (q_i) + \frac{1}{2} \sum_{\alpha, \beta = m+1}^{n} b_{\alpha\beta} (q_i) p_\alpha^0 p_\beta^0 \tag{9}$$

(see preceding section). The reduced system is acted upon by gyroscopic forces in addition to the potential forces $-\frac{\partial \Pi^*}{\partial q_i}$ $(i = 1, \ldots, m)$. Since to the stationary motion of the initial system $q_i = q_i^0$, $p_\alpha = p_\alpha^0$ $(i = 1, \ldots, m;\ \alpha = m + 1, \ldots, n)$ there corresponds an equilibrium position of the reduced system, it follows that the quantities q_i^0, p_α^0 must satisfy the equations

$$\frac{\partial \Pi^* (q_i, p_\alpha^0)}{\partial q_i} = 0 \quad (i = 1, \ldots, m) \tag{10}$$

These are necessary and sufficient conditions for the existence of stationary motion. They are obviously equivalent to the conditions (5) and are obtained from the latter by elimination of the quantities p_i $(i = 1, \ldots, m)$. Applying the Lagrange theorem to the position of equilibrium $q_i = q_i^0$ of the reduced system, we get the criterion of stability of stationary motion in the following form.

Motion with the initial conditions q_i^0, $\dot{q}_i^0 = 0$, p_α^0 $(i = 1, \ldots, m;$ $\alpha = m + 1, \ldots, n)$ *is a stable stationary motion if the Routh potential* $\Pi^* (q_i, p_\alpha^0)$ *has a strict minimum at* $q_i = q_i^0$ $(i = 1, \ldots, m)$.

By applying the Lagrange theorem to a reduced system, we fixed the constant values of the cyclic momenta $p_\alpha = p_\alpha^0$ $(\alpha = m + 1, \ldots, n)$. However, the criterion is still valid for variations of momenta p_α^0 as well. To establish this fact it is sufficient to take, for the Lyapunov function, the integral of motion

$$[E^* (q_i^0 + \xi_i, \dot{q}_i, p_\alpha^0) - E^* (q_i^0, 0, p_\alpha^0)]^2 + \sum_{\alpha = m+1}^{n} \eta_\alpha^2 \tag{11}$$

where $E^* = T^* + \Pi^*$ is the total energy of the reduced system (it coincides with the total energy of the initial system expressed in the Routh potentials q_i, $\dot{q}_i$, p_α (see Sec. 48), while $\xi_i = q_i - q_i^0$, $\dot{q}_i$, $\eta_\alpha = p_\alpha - p_\alpha^0$ $(i = 1, \ldots, m;\ \alpha = m + 1, \ldots, n)$ are deviations of perturbed motion (from the given stationary motion). The function (11) has a strict minimum (equal to zero) at $\xi_i = 0$,

$\dot{q}_i = 0, \eta_\alpha = 0$ $(i = 1, \ldots, m; \alpha = m + 1, \ldots, n)$.

The stability criterion of stationary motion given here was established in somewhat different form by Routh in 1884.

Example. Determine the stable stationary motions of a nonhomogeneous ponderable sphere on a smooth horizontal plane if the centre of gravity of the sphere, D, is distant d from the geometric centre O, the mass of the sphere is M, the moment of inertia about the OD axis is C, and the two other principal central moments of inertia are equal: $A = B$ (Fig. 61).*

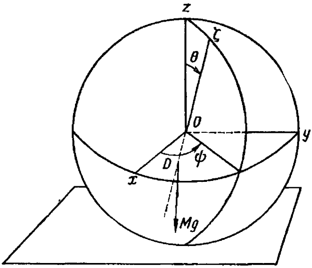

Fig. 61

For the independent coordinates, take two horizontal coordinates of the centre of gravity x_D, y_D and the three Euler angles φ, ψ, θ. Here, the angle φ is the angle of "pure rotation" about the axis of the dynamic symmetry $O\zeta$ passing through the points D and O; the direction of the $O\zeta$ axis coincides with the direction of the vector $\boldsymbol{DO}$. Write the expressions for the kinetic and potential energies:

$$2T = A(p^2 + q^2) + Cr^2 + M(\dot{x}_D^2 + \dot{y}_D^2 + \dot{z}_D^2), \quad \Pi = Mgz_D$$

Here p, q, r are projections of the angular velocity $\boldsymbol{\omega}$ on the central axes of inertia. But

$$p^2 + q^2 = \dot{\theta}^2 + \dot{\psi}^2 \sin^2\theta, \quad r = \dot{\varphi} + \dot{\psi}\cos\theta, \quad z_D = -d\cos\theta$$

Therefore

$$2T = (A + Md^2\sin^2\theta)\,\dot{\theta}^2 + A\dot{\psi}^2\sin^2\theta + Cr^2 + M(\dot{x}_D^2 + \dot{y}_D^2)$$
$$\Pi = -Mgd\cos\theta$$

The coordinates x_D, y_D, φ and ψ are cyclic. During motion, the appropriate momenta retain constant values, namely:

$$\left.\begin{array}{l} p_x = M\dot{x}_D = \alpha, \quad p_y = M\dot{y}_D = \beta, \quad p_\varphi = Cr = \gamma \\ p_\psi = A\dot{\psi}\sin^2\theta + Cr\cos\theta = \delta \end{array}\right\} \tag{12}$$

---

[*] The rigid body considered in this problem is a physical pendulum, the axis of suspension being the axis of dynamic symmetry, and the point of support O can slide freely (without friction) on the horizontal plane.

In addition, $p_\theta = (A + Md^2 \sin^2\theta)\dot{\theta}$.

Write the expression for the Hamiltonian function in the variables θ, $p_x = \alpha$, $p_y = \beta$, p_θ, $p_\varphi = \gamma$, $p_\psi = \delta$:

$$H = \frac{p_\theta^2}{2(A + Md^2\sin^2\theta)} + \frac{1}{2A}\left(\frac{\delta - \gamma\cos\theta}{\sin\theta}\right)^2 - Mgd\cos\theta + \frac{\alpha^2 + \beta^2}{2M} =$$

$$= \frac{p_\theta^2}{A + Md^2\sin^2\theta} + \Pi^* \tag{13}$$

where

$$\Pi^* = \frac{1}{2A}\left(\frac{\delta - \gamma\cos\theta}{\sin\theta}\right)^2 - Mgd\cos\theta + \frac{\alpha^2 + \beta^2}{2M} \tag{14}$$

is the Routh potential.*

Here, the conditions for the existence of stationary motion (5) are of the form $\frac{\partial\Pi^*}{\partial\theta} = 0$, $p_\theta = 0$.

The condition for stability—the presence of a strict extremum of the function H at $p_\theta = 0$ and at a certain desired value of θ—will be fulfilled if for this value of θ the function Π^* has a strict minimum. To find this value $\theta = \theta_0$, put $u = \cos\theta$ and

$$f(u) = A\Pi^* = \frac{1}{2}\frac{(\delta - \gamma u)^2}{1 - u^2} - Ku + \text{const} \quad (K = AMgd)$$

We find

$$f'(u) = \frac{-\gamma\delta u^2 + (\gamma^2 + \delta^2)u - \gamma\delta}{(1 - u^2)^2} - K$$

$$f''(u) = \frac{\gamma^2 + \delta^2}{(1 - u^2)^3}(1 - 3\mu u + 3u^2 - \mu u^3)$$

where** $\mu = \frac{2\gamma\delta}{\gamma^2 + \delta^2}$.

Suppose that the equation $f'(u) = 0$ has a root u such that $|u| < 1$. It is to this value $|u| = \cos\theta$ that the stationary motion of a sphere corresponds, for which motion the centre of the sphere is in uniform rectilinear motion and the angles φ and ψ vary by a linear law.

To find the stability of stationary motion, we shall first prove that $f''(u) > 0$ when $|u| < 1$. Indeed, if $f''(u)$ were equal to zero for $|u| < 1$, then from the expression for $f''(u)$ it would follow that $\mu = \frac{1 + 3u^2}{u^3 + 3u}$. From this it is readily seen that $|\mu| > 1$ for $|u| < 1$, which is impossible, since $\mu = \frac{2\gamma\delta}{\gamma^2 + \delta^2}$. Consequently, $f''(u) \neq 0$ when $|u| < 1$, i.e. $f''(u)$ retains its sign in the interval $(-1, +1)$. But $f''(0) = \gamma^2 + \delta^2 > 0$. Hence, $f''(u) > 0$ for $|u| < 1$. Since

$$A\frac{d^2\Pi^*}{d\theta^2} = f''(u)\sin^2\theta - f'(u)\cos\theta = f''(u)\sin^2\theta > 0$$

---

[*] We do not consider the singular values $\theta = 0$ and $\theta = \pi$. Therefore $\sin\theta \neq 0$.

[**] Here $\gamma^2 + \delta^2 > 0$ since for $\gamma = \delta = 0$ the function $\Pi^* = -Mgd\cos\theta$ has a strict minimum at $\theta = 0$.

it follows that for the value of θ under consideration the function Π^* has a strict minimum, i.e. the appropriate stationary motion is stable.

The condition for the existence of stationary motion $f'(u)=0$ may be transformed by putting $\delta=\gamma u+A\dot{\psi}(1-u^2)$.

If we then put $\gamma=Cr=C(\dot{\varphi}+\dot{\psi}\cos\theta)$ into the equation obtained, this condition will take the final form

$$C\dot{\varphi}\dot{\psi}+(C-A)\dot{\psi}^2\cos\theta+Mgd=0 \tag{15}$$

This is the well-known condition for the existence of regular precession under the action of an external moment $Mgd\sin\theta$ (the moment of the vertical reaction $N=Mg$ about the pole D).

Consider three separate cases.

1°. If $|Mgd+C\dot{\varphi}\dot{\psi}|>|A-C|\dot{\psi}^2$ then the condition (15) is not fulfilled for any real value of θ, and there does not exist any stationary motion with such angular velocities.

2°. If $|Mgd+C\dot{\varphi}\dot{\psi}|<|A-C|\dot{\psi}^2$ and the quantities $Mgd+C\dot{\varphi}\dot{\psi}$ and $A-C$ have the same sign, then for such angular velocities there exists stationary motion with $\cos\theta>0$.

3°. But if $|Mgd+C\dot{\varphi}\dot{\psi}|<|A-C|\dot{\psi}^2$ and the quantities $Mgd+C\dot{\varphi}\dot{\psi}$ and $A-C$ have different signs, then $\cos\theta<0$ in the case of stationary motion. In this case there exists stable stationary motion such that the centre of gravity is located above the geometric centre of the sphere.

Now consider special cases.

(a) $\theta_0=0$: Then from formulas (12) it follows that $\gamma=\delta$. Therefore

$$f'(u)=-\frac{\gamma^2}{(1+u)^2}-K$$

$$A\left(\frac{d\Pi^*}{d\theta}\right)_{\theta=0}=f'(u)(-\sin\theta_0)=0$$

$$A\left(\frac{d^2\Pi^*}{d\theta^2}\right)_{\theta=0}=f'(u)\,(-\cos\theta_0)=\frac{\gamma^2}{(1+u_0)^2}+K>0$$

The stationary motion is always stable.

(b) $\theta_0=\pi$. From the formulas (12) we find $\gamma=-\delta$. Therefore

$$f'(u)=\frac{\gamma^2}{(1-u)^2}-K$$

$$A\left(\frac{d\Pi^*}{d\theta}\right)_{\theta=\pi}=f'(u)(-\sin\theta_0)=0$$

$$A\left(\frac{d^2\Pi^*}{d\theta^2}\right)_{\theta=\pi}=f'(u)\,(-\cos\theta_0)=\frac{\gamma^2}{(1-u_0)^2}-K=\frac{\gamma^2}{4}-K$$

The stationary motion will be stable if the condition $\frac{\gamma^2}{4}>K$ is fulfilled, which in expanded notation comes out like this:

$$C^2r^2>4AMgd \tag{16}$$

If inequality (16) occurs, then, although in the case under consideration the centre of gravity is located over the geometric centre of the sphere, rotation about the vertical axis will be stable stationary motion.

References

1. Appell P. E. Traité de mécanique rationelle. Vols. I and II. Gauthier-Villars, Paris, 1909.
2. Babakov L. M. Theory of Oscillations. Gostekhizdat, Moscow, 1958 (in Russian).
3. Bulgakov B. V. Oscillations. Gostekhizdat, Moscow, 1954 (in Russian).
4. Cartan E. Leçons sur les invariants intégraux. Hermann, Paris, 1922.
5. Chetayev N. G. Stability of Motion, 3rd ed. "Nauka", Moscow, 1965 (in Russian).
6. Corden H. C., Stehle P. Classical Mechanics. Wiley, New York; Chapman, London, 1950.
7. Gantmacher F. R., Krein M. G. Oscillatory Matrices and Nuclei and Small Oscillations of Mechanical Systems, 2nd ed. Gostekhizdat, Moscow, 1950 (in Russian).
8. Goldstein H. Classical Mechanics. Addison-Wesley, Reading, Massachusetts, 1951.
9. Jacobi K.G. J. Vorlesungen über Dynamik. Clebsch, 1866.
10. Lagrange J. L. Mécanique analytique. Paris, 1788.
11. Lanczos C. The Variational Principles of Mechanics. University of Toronto Press, Toronto, 1949.
12. Landau L. D., Lifshits E. M. Mechanics. Fizmatgiz, 1958 (in Russian).
13. La Salle J., Lefschetz S. Stability by Lyapunov's Direct Method with Applications. Academic Press, New York and London, 1961.
14. La Vallée-Poussin Ch. J. de, Leçons de mécanique analytique, tome 1-2. Paris, 1925.
15. Loitsyansky L. G., Lurie A. I. Theoretical Mechanics, Part III. ONTI. Moscow, 1934 (in Russian).
16. Lyapunov A. M. The General Problem of Stability of Motion. Gostekhizdat, Moscow, 1950 (in Russian).
17. MacMillan W. D. Dynamics of Rigid Bodies. McGraw-Hill, New York and London, 1936.
18. Malkin I. G. Theory of Stability of Motion. Gostekhizdat, Moscow, 1952 (in Russian).
19. Merkin D. R. Gyroscopic Systems. Gostekhizdat, Moscow, 1956 (in Russian).
20. Rose N. V. Lectures in Analytical Mechanics, Part I. Leningrad University Press, Leningrad, 1938 (in Russian).
21. Routh E. T. The Advanced Part of a Treatise on the Dynamics of a System of Rigid Bodies, 6th ed. Macmillan, London, 1905.
22. Siegel C. L. Vorlesungen über Himmelsmechanik. Springer, Berlin, 1956.
23. Sommerfeld A. Mechanik. 4te neubearb. Aufl. Geest und Portig, Leipzig, 1948.
24. Suslov G. K. The Principles of Analytical Mechanics, 2nd ed. Kiev, 1911-1912; 3rd ed. revised by N. N. Buchholtz and V. K. Goltzman, entitled: Theoretical Mechanics. Gostekhizdat, Moscow, 1944 (in Russian).
25. Synge J. L. Classical Dynamics. McGraw-Hill, New York, 1958.
26. Variational Principles. Collection of articles, ed. by L. S. Polak. Fizmatgiz, Moscow, 1959 (in Russian).
27. Whittaker E. T. A Treatise on the Analytical Dynamics of Particles and Rigid Bodies. Dover Publications, New York. 1944.

Name Index

Aristotle 25

Bernoulli, J. 25

Carathéodory, C. 154
Cartan, E. 8, 119
Chetayev, N. G. 173, 176
Courant, R. 219

Dirichlet, P. G. Lejeune 169

Fischer, E. 219
Fermat, P. 115

Galileo, G. 25, 28

Hamilton, W. R. 73, 89
Helmholtz, H. L. F. 105
Hertz, H. 248
Hurwitz, A. 198
Huyghens, C. 28
Hwa-Chung, Lee 122

Jacobi, K. G. J. 84, 114

Lagrange, J. L. 22, 115, 169, 210
Landau, L. 86
Lyapunov, A. M. 173, 176, 181, 193

Malkin, I. G. 176
Maupertuis, P. 115
Mikhailov, A. V. 200

Ostrogradsky, M. V. 89

Poincaré, J. H. 8, 119, 193
Poisson, S. D. 84
Pyatigorsky, L. 86

Rayleigh, J. W. 220
Rose, H. V. 115
Routh, E. 198, 243, 245, 254, 256

Stoker, J. J. 194
Suslov, G. K. 7, 114

Tait, P. 248
Thomson, W. 105, 248
Torricelli, E. 169

Vallée-Poussin, Ch. J. 7

Webster, A. G. 248
Weierstrass, K. 210
Whittaker, E. T. 112

Subject Index

Accelerations, generalized 41
Action, elementary 88
  Hamilton 88, 115
  Lagrange 114, 115
Affix (of a complex number) 200
Analytical mechanics 7, 8
Angle, of nutation 35
  of precession 35
  of pure rotation 35
Appell's equations 22, 61
  for nonholonomic systems 57 ff.
Associated expression (of a function) 72, 73
Asymptotic stability, of equilibrium position 176 ff.
  of linear system, criteria of 197
  theorem on 178
Autonomous system (of differential equations) 81

Bolzano-Weierstrass lemma 179

Canonical equations 73
  of Hamilton 71 ff.
Canonical transformation(s) 128 ff.
  free 132 ff., 135
  nonfree 132, 133
  nonpoint 135
  structure of an arbitrary 152 ff.
  univalent 131
  valence of 131
Canonicity of transformation 163, 164
Characteristic equation 190
Characteristic numbers 190
Characteristic, amplitude-phase 236, 237
  frequency 236, 239
Chetayev's instability test 186
Chetayev's theorem 176
Circulation of velocity 107
Compatibility with constraints 25
Complete integral 136
Complex spectrum 238
Conditional stability 182 ff.
Conjugate kinetic focus 95
Conservation of energy, law of 77
Constrained systems 9
Constraint(s) 9
  bilateral 12
  classification of 9 ff.
  differential 9
  differential stationary 11
  finite 9
  finite, integrable 10
  finite, stationary 11
  "frozen" 14
  geometric 9
  ideal 17 ff.
  kinematical 9
  nonholonomic 11
  nonstationary 18
  semiholonomic 11
  taut 12
  unilateral 12
Contact transformations 131
Coordinates, cyclic 80 ff.
  ignorable 80 ff.
  kinosthenic 82
  noncyclic 80, 243
  normal 174, 210 ff.
  position 243
  principal 174, 212
Coriolis force 52
Coupled pendulums 210
Curl (of velocity) 107
Cyclic coordinates 80 ff.

D'Alembert's principle 30, 31
Degree of stability 192
Degrees of freedom 15
Deviations 166
Dirichlet's theorem 169
Displaced lines 110

- Displacements, possible 13, 14
  - possible infinitesimal 13
  - virtual 13, 14
- Dissipation of energy, total 53, 54
- Dissipative forces 51 ff.
- Dissipative system 52
  - definite 54
- Donkin's theorem 74, 75
- Dynamical equations of Euler 64, 186

- Effective forces 16
- Eigenvector 191
- Elementary action 88
- Energy integral 51
- Equation, characteristic 190
  - frequency 205
  - secular 190, 205, 230
- Equilibrium position 25, 166
  - asymptotically stable 176
  - stable 166, 167
  - unstable 167
- Eulerian angles 35
- Euler's differential equations 91
- Euler's dynamical equations 64, 186
- Euler's equations 64, 65
- Euler's formula 49
- Euler's theorem on homogeneous functions 76

- Fermat's principle 115, 116
- Filter 238
- Fluid line 107
- Fluid surface 107
- Free body 19
- Free systems 9
- Frequency equation 205

- General equation of dynamics 20 ff., 30
- Generalized accelerations 41
- Generalized conservative system(s) 76, 111 ff.
- Generalized coordinates, independent 34
- Generalized forces 60
- Generalized integral of energy 76
- Generalized momenta 71
- Generalized potential 67, 68
- Generalized total energy 76
- Generalized velocities 41, 68
- Generating function 131
- Geodesic lines 117, 118
- Golden rule of mechanics 25, 26
- Group 128
- Gyroscopic forces 51
- Gyroscopic moment 53
- Gyroscopically uncoupled system 246 ff.

- Hamilton action 115
- Hamiltonian function 73, 77, 80
- Hamiltonian variables 71, 72, 75, 77
- Hamilton-Jacobi equation 135 ff., 143
- Hamilton-Ostrogradsky principle 89
- Hamilton's canonical equations 102, 104, 106
- Hamilton's equations 73
- Hamilton's principal function 137-139
- Hamilton's principle 88 ff.
  - second form of 96 ff.
- Helmholtz theorem 110
- Hessian 70, 71, 74, 75
  - small 78
- Hidden motions 248
- Holonomic rheonomic system 12
- Holonomic scleronomic system 11
- Holonomic system 11, 33 ff.
- Hurwitz determinant 198, 199
- Hurwitz polynomial 201
- Hydrodynamical interpretation of integral invariant of mechanics 105 ff.
- Hypothesis of realization of admissible reaction forces 21

- Ignorable coordinates 80 ff.
- Influence coefficient 224
- Integral of equations of motion 83
- Integral invariant(s) 8, 101
  - absolute 122
  - in extended coordinate space 107
  - of mechanics, basic 98 ff.
  - relative 121
  - universal 122
  - universal relative 122
- Invariance of Poincaré integral, geometrical interpretation of 120
- Invariance of volume in phase space 125 ff.

- Jacobi equations 114
- Jacobi theorem 136
- Jacobian 71, 74
- Jacobian functional matrix 47
- Jacobian matrix, simplicial nature of 161

of order n 161
Jacobi-Poisson theorem 85

Keplerian motion 146, 147
Kepler's three laws of planetary motion 148
Kinetic potential 66
Kinosthenic coordinates 82
Kronecker's symbol 160, 189

Lagrange action 114, 115
Lagrange brackets 160
Lagrange's equations, investigations of 45 ff.
Lagrange's equations for potential forces 66 ff.
Lagrange's equations in independent coordinates 41
Lagrange's equations of the first kind 22
Lagrange's equations of the second kind 22, 63
  in independent coordinates 40 ff.
Lagrange's theorem 169, 171
Lagrange's theorem on stability of equilibrium position 166 ff.
Lagrangian function 66, 73, 245
Lagrangian variables 71, 72, 75, 78
Legendre transformation 74
Liénard-Chipart conditions 199
Liénard-Chipart criterion 201
Linear approximation for nonlinear system 193
Liouville's theorem 127, 150
Lorentz force 68
Lyapunov function 182, 254
Lyapunov's first theorem 174
Lyapunov's second theorem 176
Lyapunov's theorem 184, 185
  corollary to 184, 185

Matrix, oscillatory 225
  generalized-simplicial 163
  positive definite 205
  simplicial 162
Maupertuis-Lagrange principle of least action 114 ff.
Maximinimal properties 219
Maxwell's principle of reciprocity 224
Mersenne law 227
Mikhailov's criterion 200
Minimaximal properties 219

Natural systems 70
Newton's law of gravitation 148
Nodal line 35
Noncyclic coordinates 80, 243
Nonholonomic system 11, 12
Nonnatural system 70
Normal coordinates 174, 210 ff.

Optical principle of Fermat 115, 116
Oscillations, forced 216, 236
  free 216
Oscillatory matrix 225

Paths, circuitous 89
  straight-line 89
Perturbation theory, application of canonical transformations to 150 ff.
Poincaré integral 121
Poincaré-Cartan integral 101, 104, 106, 121, 129
Poincaré-Cartan integral invariant 98 ff.
Point transformation 135
Poisson bracket(s) 83 ff., 160
  invariance of, in canonical transformation 163 ff.
Poisson identity 84
Position coordinates 243
Potential energy 48
Potential forces 48
Potential of forces 48
Power of nonpotential forces 50
Principal coordinates 174, 212
Principal oscillations 210
Principle of least action of Maupertuis-Lagrange 114, 115
Principle of possible displacements 25, 29
Principle of reciprocity, Maxwell's 224
Principle of virtual displacements 25, 26, 30, 39
Pseudoaccelerations 60, 62
Pseudocoordinates 59
Pseudovelocities 59

Rayleigh's dissipative function 53, 231
Rayleigh's functions 233
Rayleigh's theorem 220
Reaction forces of constraints 16
Reduced system 246
Resonance 216

Rheonomic system 11
Rotor (of velocity) 107
Routh equations 78 ff., 80, 81
Routh functions 80, 81, 244
Routh potential 245, 246, 252, 257
Routh variables 78, 242, 243
Routh-Hurwitz condition 199

Scleronomic system 11, 35
Secular equation 190, 205, 230
Secular terms 192, 209
Separation of variables, method of 141 ff.
Separation theorem 221
Shortest lines in space 118
Simplicial matrix 162
Simplicial group 162
Solution, asymptotically stable 185
  stable 185
Spectrum, complex 238 Ш
Stability, geometric criterion of 200
  of linear systems 189 ff.
  of stationary motions 252 ff.
Stationary motion, conditions for existence of 255
  criterion of stability of 254, 256
  stability of 252 ff.
  stable 254, 255
Statistical ensemble 127
Theorem of Lee Hwa-Chung 122, 130
Theorem on asymptotic stability 178
Theorem on conservation of circulation of velocity 108
Theorem on conservation of vortex flow 108
Theorem on stability of a dissipative system 181
Thomson's theorem 107
Torricelli's principle 27
Total energy 170
Transformation (*see* canonical point)
  testing canonical character of 158 ff.

Undetermined multipliers of Lagrange 21, 22
Universal integral invariant 121, 130
  of Poincaré 119 ff.

Valence of canonical transformation 131, 161
Velocities, generalized 41, 68
  possible 13
Vortex lines 109, 110

Whittaker's equations 112

Printed in the Union of Soviet Socialist Republics